THE ANNALS *of* MATHEMATICAL STATISTICS

(FOUNDED BY H. C. CARVER)

THE OFFICIAL JOURNAL OF THE INSTITUTE OF MATHEMATICAL STATISTICS

Contents

Vol. XII, No. 4 — December, 1941

THE ANNALS OF MATHEMATICAL STATISTICS

The ANNALS OF MATHEMATICAL STATISTICS is published quarterly by the Institute of Mathematical Statistics, Mt. Royal & Guilford Aves., Baltimore, Md. Subscriptions, renewals, orders for back numbers and other business communications should be sent to the ANNALS OF MATHEMATICAL STATISTICS, Mt. Royal & Guilford Aves., Baltimore, Md., or to the Secretary of the Institute of Mathematical Statistics, E. G. Olds, Carnegie Institute of Technology, Pittsburgh, Pa. Changes in mailing address which are to become effective for a given issue should be reported to the Secretary on or before the 15th of the month preceding the month of that issue. The months of issue are March, June, September and December.

Manuscripts for publication in the ANNALS OF MATHEMATICAL STATISTICS should be sent to S. S. Wilks, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten double-spaced with wide margins, and the original copy should be submitted. Footnotes should be reduced to a minimum and whenever possible replaced by a bibliography at the end of the paper; formulae in footnotes should be avoided. Figures, charts, and diagrams should be drawn on plain white paper or tracing cloth in black India ink twice the size they are to be printed. Authors are requested to keep in mind typographical difficulties of complicated mathematical formulae.

Authors will ordinarily receive only galley proofs. Fifty reprints without covers will be furnished free. Additional reprints and covers furnished at cost.

The subscription price for the ANNALS is $4.00 per year. Single copies $1.25. Back numbers are available at the following rates:

Vols. I-IV $5.00 each. Single numbers $1.50.
Vols. V to date $4.00 each. Single numbers $1.25.

COMPOSED AND PRINTED AT THE
WAVERLY PRESS, INC.
BALTIMORE, MD., U. S. A.

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879

the
nore,
com-
Mt.
nsti-
logy,
e for
f the
arch,

STICS
ripts
copy
never
foot-
n on
re to
ulties

thout
cost.

$1.25.

1879

DISTRIBUTION OF THE RATIO OF THE MEAN SQUARE SUCCESSIVE DIFFERENCE TO THE VARIANCE

By John von Neumann

Institute for Advanced Study[1]

1. Introduction. Let $x_1, \cdots, x_n$ be variables representing n successive observations in a population which obeys a distribution law

$$ce^{-(x-\xi)^2/2\sigma^2}\,dx, \qquad \left(c = \frac{1}{\sigma\sqrt{2\pi}}\right),$$

i.e. which is normal, with the mean ξ and the standard deviation σ. For the sample we define as usual the mean,

$$\bar{x} = \frac{1}{n}\sum_{\mu=1}^{n} x_\mu,$$

the variance,

$$s^2 = \frac{1}{n}\sum_{\mu=1}^{n} (x_\mu - \bar{x})^2,$$

and also the mean square successive difference

$$\delta^2 = \frac{1}{n-1}\sum_{\mu=1}^{n-1} (x_{\mu+1} - x_\mu)^2.$$

The reasons for the study of the distribution of the mean square successive difference δ^2, in itself as well as in its relationship to the variance s^2, have been set forth in a previous publication[2], to which the reader is referred. The distribution of δ^2, and in particular its moments, were also studied there. The present paper is devoted to the investigation of the ratio

$$\eta = \frac{\delta^2}{s^2}.$$

A comparison of the observed value of η with that distribution is particularly suited as a basis of the judgment whether the observations $x_1, \cdots, x_n$ are independent or whether a trend exists. (Cf. sections 1 and 2, loc. cit.[2])

The moments of η have already been determined by J. D. Williams by a

---

[1] Also Scientific Advisory Committee of the Ballistic Research Laboratory, Aberdeen Proving Ground.

[2] John von Neumann, R. H. Kent, H. R. Bellinson, B. I. Hart, "The mean square successive difference," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 153–162.

different method.[3] Williams' results have been checked by W. J. Dixon at the suggestion of S. S. Wilks, whose stimulating interest has been largely responsible for the undertaking of the series of papers on δ^2 and $\frac{\delta^2}{s^2}$. The present rather exhaustive discussion, however, brings out several other essential characteristics of this statistic, and provides the key to some very effective computational methods. It is further hoped that the reader will find that the mathematical methods used and the generalizations indicated have an interest of their own.

From the latter point of view the final results of sections 5 and 7, concerning the distribution of values of quadratic and of Hermitian forms, may deserve special attention.

2. Diagonalization of the quadratic forms and replacement by a spherical mean. Since δ^2 and s^2 are unchanged when we replace each x_μ by $x_\mu - \xi$, we may assume $\xi = 0$. Then the distribution law of x is

$$ce^{-x^2/2\sigma^2}\,dx, \qquad \text{and that of } x_1, \cdots, x_n \text{ is } \prod_{\mu=1}^{n} ce^{-x_\mu^2/2\sigma^2}\,dx_\mu,$$

i.e.

$$c^n e^{-\sum_{\mu=1}^{n} x_\mu^2/2\sigma^2}\,dx_1 \cdots dx_n.$$

Any linear orthogonal transformation of the $x_1, \cdots, x_n$ leaves $\sum_{\mu=1}^{n} x_\mu^2$ and $dx_1 \cdots dx_n$ unchanged, hence the above distribution law will likewise be left unchanged. Thus, we may subject the two quadratic forms δ^2, s^2 to any simultaneous linear, orthogonal transformation.

Consider one carrying $x_1, \cdots, x_n$ into, say $x_1', \cdots, x_n'$, which brings the quadratic form $(n-1)\delta^2$ into the diagonal form, say $\sum_{\mu=1}^{n} A_\mu x_\mu'^2$. Such a transformation does not affect the characteristic values of the quadratic forms[4], and these characteristic values are obviously $A_1, \cdots, A_n$ in the case of $\sum_{\mu=1}^{n} A_\mu x_\mu'^2$. Consequently $A_1, \cdots, A_n$ are the characteristic values of the original quadratic form $(n-1)\delta^2$. We shall determine them as such in the next section.

Clearly we always have $(n-1)\delta^2 \geqq 0$, hence all $A_\mu \geqq 0$. Some A_μ may

---

[3] J. D. Williams, "Moments of the ratio of the mean square successive difference to the mean square difference in samples from a normal universe," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 239–241. Cf. also L. C. Young, "On randomness in orded sequences," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 293–300.

[4] For the properties of matrices and quadratic forms cf. e.g.: J. H. M. Wedderburn, Lectures on Matrices, *Amer. Math. Soc. Colloquium Publications*, Vol. 17, New York, 1934. In the present context cf. mainly Chapters II and VI.

equal 0 say k ($= 0, 1, \cdots, n$) of them, which we can arrange to be $A_{n-k+1}, \cdots, A_n$.

$(n-1)\delta^2 = 0$ is thus equivalent to $x'_1 = \cdots = x'_{n-k} = 0$, i.e. to $n - k$ independent conditions. On the other hand this amounts obviously to $x_1 = \cdots = x_n$, and these are $n - 1$ independent conditions. So $k = 1$ and consequently $A_1, \cdots, A_{n-1} > 0$, $A_n = 0$. And our linear orthogonal transformation must carry the x-vectors with $x_1 = \cdots = x_n$ into the x'-vectors with $x'_1 = \cdots = x'_{n-1} = 0$. Among the former, $\frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}}$ has the length 1; among the latter only $0, \cdots, 0, \pm 1$ have. Hence these correspond to each other. Now the scalar (inner) product of two vectors is an orthogonal invariant, that of a vector $x_1, \cdots, x_n$ with $\frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}}$ is $\sqrt{n}\bar{x}$, that of a vector $x'_1, \cdots, x'_n$ with $0, \cdots, 0, \pm 1$ is $\pm x'_n$, hence

$$\sqrt{n}\,\bar{x} = \pm x'_n.$$

Put $x_\mu = \bar{x} + u_\mu$. Then clearly $\sum_{\mu=1}^{n} u_\mu = 0$. Hence

$$\sum_{\mu=1}^{n} x_\mu^2 = n\bar{x}^2 + \sum_{\mu=1}^{n} u_\mu^2 = x_n'^2 + ns^2.$$

Owing to the orthogonality, the left-hand side is equal to $\sum_{\mu=1}^{n} x_\mu'^2$, therefore

$$ns^2 = \sum_{\mu=1}^{n-1} x_\mu'^2.$$

Remembering that $A_n = 0$, we also have

$$(n-1)\delta^2 = \sum_{\mu=1}^{n-1} A_\mu x_\mu'^2.$$

Consequently

$$\eta = \frac{\delta^2}{s^2} = \frac{n}{n-1} \frac{\sum_{\mu=1}^{n-1} A_\mu x_\mu'^2}{\sum_{\mu=1}^{n-1} x_\mu'^2}.$$

The distribution law is, as we know, the same in $x'_1, \cdots, x'_n$ as in $x_1, \cdots, x_n$, namely

$$c^n e^{-\sum_{\mu=1}^{n} x_\mu'^2/2\sigma^2}\, dx'_1 \cdots dx'_n.$$

Thus $x'_1, \cdots, x'_n$ are independent. η depends on $x'_1, \cdots, x'_{n-1}$ only, hence we may disregard x'_n altogether, and use the distribution law of the $x'_1, \cdots, x'_{n-1}$,

$$c^{n-1} e^{-\sum_{\mu=1}^{n-1} x_\mu'^2/2\sigma^2}\, dx'_1 \cdots dx'_{n-1}.$$

With respect to $x'_1, \cdots, x'_{n-1}$ we may now state that the $x'_1, \cdots, x'_{n-1}$ distribution of η can be obtained by determining first the distribution of η over every spherical surface

$$\sum_{\mu=1}^{n-1} x'^2_\mu = r^2$$

and then averaging these distributions with the weights $\psi(r)\, dr$, where $\psi(r)\, dr$ is the probability of the spherical shell from r to $r + dr$ with respect to our original $x'_1, \cdots, x'_{n-1}$ distribution law. (It happens to be $c'e^{-r^2/2\sigma^2}r^{n-2}\, dr$, but this is immaterial.)

Since the $x'_1, \cdots, x'_{n-1}$ distribution law is obviously spherically symmetric in these variables, the first-mentioned distributions over the spherical surfaces are readily obtained by assigning each piece of the surfaces in question its own relative, $n - 2$-dimensional area as weight.

Since η is a homogeneous function of $x'_1, \cdots, x'_{n-1}$ of order zero, these spherical surface distributions of η are the same for all r. Consequently we can replace all these r by, say $r = 1$, and the subsequent averaging over the r may be omitted altogether.

Finally, since we restrict ourselves to $r = 1$, i.e. to the spherical surface

$$\sum_{\mu=1}^{n-1} x^2_\mu = 1,$$

the denominator of η may be omitted and we have

$$\eta = \frac{n}{n-1} \sum_{\mu=1}^{n-1} A_\mu x'^2_\mu.$$

We sum up, writing again $x_1, \cdots, x_{n-1}$ for $x'_1, \cdots, x'_{n-1}$, then the desired distribution of η is that of

$$\eta = \frac{n}{n-1} \sum_{\mu=1}^{n-1} A_\mu x^2_\mu,$$

where the point $x_1, \cdots, x_{n-1}$ is uniformly distributed over the spherical surface

$$\sum_{\mu=1}^{n-1} x^2_\mu = 1.$$

Here $A_1, \cdots, A_{n-1}$ are all positive, and together with 0 they are the characteristic values of the quadratic form

$$\begin{aligned}(n-1)\delta^2 &= \sum_{\mu=1}^{n-1} (x_{\mu+1} - x_\mu)^2 \\ &= x_1^2 + 2\sum_{\mu=2}^{n-1} x^2_\mu + x^2_n - 2\sum_{\mu=1}^{n-1} x_\mu x_{\mu+1}.\end{aligned}$$

3. The characteristic values A_μ; first orientation concerning η. We have shown that there exist (counting multiplicities) precisely $n - 1$ positive roots A of the characteristic equation

$$Det \begin{vmatrix} A-1 & 1 & & & & & \\ 1 & A-2 & 1 & & & & \\ & 1 & A-2 & 1 & & & \\ & & 1 & \ddots & & & \\ & & & \ddots & 1 & & \\ & & & 1 & A-2 & 1 & \\ & & & & 1 & A-2 & 1 \\ & & & & & 1 & A-1 \end{vmatrix} = 0$$

(the empty places are filled with zeros), and that these roots are the $A_1, \cdots, A_{n-1}$.

Such an A is characterized by the possibility of solving the equations

$$(A-1)x_1 + x_2 = 0, \quad x_1 + (A-2)x_2 + x_3 = 0, \quad x_2 + (A-2)x_3 + x_4 = 0,$$
$$\cdots\cdots\cdots, \quad x_{n-2} + (\mathrm{A} - 2)x_{n-1} + x_n = 0, \quad x_{n-1} + (A-1)x_n = 0,$$

in $x_1, \cdots, x_n$ not all equal to zero. Put

$$x_0 = x_1, \qquad x_{n+1} = x_n,$$

and

$$A = 2 - 2\cos\alpha,$$

then these equations become

$$x_{\mu-1} + x_{\mu+1} = 2\cos\alpha \cdot x_\mu \qquad \text{for} \qquad \mu = 1, 2, \cdots, n-1, n.$$

The last equation is satisfied by

$$x_\mu = 2\cos(\mu - \tfrac{1}{2})\alpha \qquad \text{for} \qquad \mu = 0, 1, 2, \cdots, n-1, n, n+1.$$

Now $x_0 = x_1$ is automatically fulfilled, while $x_{n+1} = x_n$ demands $\cos(n + \frac{1}{2})\alpha = \cos(n - \frac{1}{2})\alpha$. This is certainly the case when $(n + \frac{1}{2})\alpha = 2k\pi - (n - \frac{1}{2})\alpha$ (k any integer), i.e. $\alpha = \dfrac{k\pi}{n}$. For no $k = 1, \cdots, n-1$ are $x_1, \cdots, x_n$ all equal to zero $\left(\text{indeed } x_1 = 2\cos\dfrac{k\pi}{2n} > 0\right)$, hence these k give A's of the desired kind. They are

$$A = 2 - 2\cos\frac{k\pi}{n} = 4\sin^2\frac{k\pi}{2n} \qquad (k = 1, \cdots, n-1),$$

and so they are all positive and different from each other. Their number is $n - 1$. Hence they are precisely $A_1, \cdots, A_{n-1}$.

So we have shown

$$A_\mu = 2 - 2\cos\frac{\mu\pi}{n} = 4\sin^2\frac{\mu\pi}{2n} \qquad (\mu = 1, \cdots, n-1).$$

We can now reformulate the final result of the preceding section. Let us set

$$\eta = \frac{2n}{n-1}(1-\epsilon).$$

Then

$$\epsilon = \sum_{\mu=1}^{n-1} \cos\frac{\mu\pi}{n}\cdot x_\mu^2,$$

where the point $x_1, \cdots, x_{n-1}$ is uniformly distributed over the spherical surface

$$\sum_{\mu=1}^{n-1} x_\mu^2 = 1.$$

Replacement of x_μ by $x_{n-\mu}$ carries ϵ into $-\epsilon$. Therefore the distribution of ϵ is symmetric around 0. Hence the mean of ϵ is 0. The maximum of ϵ's distribution is clearly $\cos\frac{\pi}{n}$, its minimum is $\cos\frac{(n-1)\pi}{n} = -\cos\frac{\pi}{n}$. We state these facts, together with their equivalents for η.

ϵ (η)'s distribution is symmetric around its mean, which is $0\left(\frac{2n}{n-1}\right)$. The maximum of ϵ (η)'s distribution is $\cos\frac{\pi}{n}\left(\frac{2n}{n-1}\left[1+\cos\frac{\pi}{n}\right] = \frac{4n}{n-1}\cos^2\frac{\pi}{2n}\right)$, its minimum is $-\cos\frac{\pi}{n}\left(\frac{2n}{n-1}\left[1-\cos\frac{\pi}{n}\right] = \frac{4n}{n-1}\sin^2\frac{\pi}{2n}\right)$.

Thus it will be easier to obtain information concerning η by considering the distribution of ϵ, since all odd moments of ϵ are zero, etc. The investigation of ϵ instead of η was first suggested by B. I. Hart, who also found, that the first four odd moments of ϵ vanish. R. H. Kent and B. I. Hart also determined the minima and maxima of these distributions for certain small values of n.

4. Direct computation of the moments. We shall investigate the distribution law of a quantity

$$\gamma = \sum_{\mu=1}^{m} B_\mu x_\mu^2.$$

where the point $x_1, \cdots, x_m$ is equidistributed over the spherical surface

$$\sum_{\mu=1}^{m} x_\mu^2 = 1.$$

$\left(\text{Our above } \epsilon \text{ obtains by putting } m = n-1 \text{ and } B_\mu = \cos\frac{\mu\pi}{n}.\right)$

We denote the mean of any function

$$f(x_1, \cdots, x_m)$$

over the above-mentioned spherical surface (the $x_1, \cdots, x_m$ being equidistributed over it) by

$$\overline{f(x_1, \cdots, x_m)}.$$

Our primary objective is to determine the moments of this distribution

$$M_p = \overline{\gamma^p} = \overline{\left(\sum_{\mu=1}^{m} B_\mu x_\mu^2\right)^p}, \qquad (p = 0, 1, 2, \cdots).$$

Let us write Σ_m for the (m − 1-dimensional) area of the above-mentioned spherical surface (of the unit sphere in m-dimensional Euclidean space).

Now we form the function

$$f(z) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{z \sum_{\mu=1}^{m} B_\mu x_\mu^2} e^{-\sum_{\mu=1}^{m} x_\mu^2} \cdot dx_1 \cdots dx_m .$$

(This integral, as well as all others which we are going to derive from it, is obviously convergent, as long as z is sufficiently small. More precisely this is true when

$$| z | \cdot \operatorname{Max} (| B_1 |, \cdots, | B_m |) \leqq 1.$$

We shall use them only in the neighborhood of $z = 0$.) Now clearly

$$\begin{aligned}
\left\{\frac{d^p}{dz^p} f(z)\right\}_{z=0} &= \left\{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\sum_{\mu=1}^{m} B_\mu x_\mu^2\right)^p e^{z \sum_{\mu=1}^{m} B_\mu x_\mu^2} e^{-\sum_{\mu=1}^{m} x_\mu^2} dx_1 \cdots dx_m\right\}_{z=0} \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\sum_{\mu=1}^{m} B_\mu x_\mu^2\right)^p e^{-\sum_{\mu=1}^{m} x_\mu^2} dx_1 \cdots dx_m \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} M_p \left(\sum_{\mu=1}^{m} x_\mu^2\right)^p e^{-\sum_{\mu=1}^{m} x_\mu^2} dx_1 \cdots dx_m \\
&= \int_0^{\infty} M_p r^{2p} e^{-r^2} \Sigma_m r^{m-1} dr \\
&= \Sigma_m M_p \int_0^{\infty} e^{-r^2} r^{2p+m-1} dr \text{ [5]} \\
&= \tfrac{1}{2} \Sigma_m M_p \int_0^{\infty} e^{-u} u^{p+\frac{1}{2}m-1} du \\
&= \tfrac{1}{2} \Sigma_m M_p \Gamma\left(p + \frac{m}{2}\right).
\end{aligned}$$

[5] Introduce the new integration variable $u = r^2$.

On the other hand

$$f(z) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\sum_{\mu=1}^{m} (1-B_\mu z)x_\mu^2}\, dx_1 \cdots dx_m$$

$$= \prod_{\mu=1}^{m} \int_{-\infty}^{\infty} e^{-(1-B_\mu z)x_\mu^2}\, dx_\mu \quad {}^{6}$$

$$= \prod_{\mu=1}^{m} \tfrac{1}{2}(1 - B_\mu z)^{-\frac{1}{2}} \cdot 2 \int_0^\infty e^{-u} u^{-\frac{1}{2}}\, du$$

$$= \prod_{\mu=1}^{m} \tfrac{1}{2}(1 - B_\mu z)^{-\frac{1}{2}} \cdot 2\Gamma(\tfrac{1}{2})$$

$$= \Gamma(\tfrac{1}{2})^m \mathfrak{P}(z)^{-\frac{1}{2}},$$

where

$$\mathfrak{P}(z) = \prod_{\mu=1}^{m} (1 - B_\mu z).$$

Thus

$$\tfrac{1}{2}\Sigma_m M_p \Gamma\left(p + \frac{m}{2}\right) = \Gamma\left(\frac{1}{2}\right)^m \left\{\frac{d^p}{dz^p} \mathfrak{P}(z)^{-\frac{1}{2}}\right\}_{z=0}.$$

For $p = 0$ this becomes, since $M_0 = 1$, $\mathfrak{P}(0) = 1$,

$$\frac{1}{2} \Sigma_m \Gamma\left(\frac{m}{2}\right) = \Gamma\left(\frac{1}{2}\right)^m.$$

Dividing the former equation by the latter gives, since

$$\frac{\Gamma\left(p + \frac{m}{2}\right)}{\Gamma\left(\frac{m}{2}\right)} = \frac{m}{2}\left(\frac{m}{2} + 1\right) \cdots \left(\frac{m}{2} + p - 1\right),$$

$$M_p = \frac{1}{\frac{m}{2}\left(\frac{m}{2} + 1\right) \cdots \left(\frac{m}{2} + p - 1\right)} \left\{\frac{d^p}{dz^p} \mathfrak{P}(z)^{-\frac{1}{2}}\right\}_{z=0}.$$

In order to make a practical use of the above formula, we compute

$$\ln\left(\mathfrak{P}(z)^{-\frac{1}{2}}\right) = -\tfrac{1}{2} \sum_{\mu=1}^{m} \ln(1 - B_\mu z)$$

$$= -\tfrac{1}{2} \sum_{\mu=1}^{m} \sum_{l=1}^{\infty} -\frac{1}{l} B_\mu^l z^l$$

$$= \sum_{l=1}^{\infty} \frac{1}{2l} \left(\sum_{\mu=1}^{m} B_\mu^l\right) z^l.$$

[6] Introduce the new integration variable $u = (1 - B_\mu z)r^2$.

Write

$$\alpha_l = \frac{1}{2l}\sum_{\mu=1}^{m} B_\mu^l,$$

then

$$\mathfrak{P}(z)^{-\frac{1}{2}} = e^{\alpha_1 z + \alpha_2 z^2 + \alpha_3 z^3 + \cdots}$$

$$= 1 + \beta_1 z + \beta_2 z^2 + \beta_3 z^3 + \cdots,$$

and so

$$M_p = \frac{1\cdot 2\cdots p}{\frac{m}{2}\left(\frac{m}{2}+1\right)\cdots\left(\frac{m}{2}+p-1\right)}\beta_p .$$

Clearly

$$\beta_1 = \alpha_1 ,$$

$$\beta_2 = \alpha_2 + \tfrac{1}{2}\alpha_1^2 ,$$

$$\beta_3 = \alpha_3 + \alpha_1\alpha_2 + \tfrac{1}{6}\alpha_1^3 ,$$

$$\beta_4 = \alpha_4 + \tfrac{1}{2}\alpha_2^2 + \alpha_1\alpha_3 + \tfrac{1}{2}\alpha_1^2\alpha_2 + \tfrac{1}{24}\alpha_1^4 .$$

In our application (cf. above)

$$B_{m+1-\mu} = -B_\mu .$$

This has the consequence that

$$\alpha_1 = \alpha_3 = \alpha_5 = \cdots = 0.$$

Thus the z functions we compute contain only even powers of z and consequently

$$\beta_1 = \beta_3 = \beta_5 = \cdots = 0,$$

$$M_1 = M_3 = M_5 = \cdots = 0,$$

and

$$\beta_2 = \alpha_2 ,$$

$$\beta_4 = \alpha_4 + \tfrac{1}{2}\alpha_2^2 ,$$

$$\beta_6 = \alpha_6 + \alpha_2\alpha_4 + \tfrac{1}{6}\alpha_2^3 ,$$

$$\beta_8 = \alpha_8 + \tfrac{1}{2}\alpha_4^2 + \alpha_2\alpha_6 + \tfrac{1}{2}\alpha_2^2\alpha_4 + \tfrac{1}{24}\alpha_2^4 .$$

As mentioned before, we actually have $m = n - 1$ and $B_\mu \equiv \cos \dfrac{\mu\pi}{n}$. Consequently

$$\alpha_l = \frac{1}{2l} \sum_{\mu=1}^{n-1} \left\{\cos \frac{\mu\pi}{n}\right\}^l = \frac{1}{2l} \sum_{\mu=1}^{n-1} \{\tfrac{1}{2}(e^{i\mu\pi/n} + e^{-i\mu\pi/n})\}^l$$

$$= \frac{1}{2^{l+1}l} \sum_{\mu=1}^{n-1} \sum_{k=0}^{l} \binom{l}{k} e^{i(2k-l)\mu\pi/n} \text{ [7]}$$

$$= \frac{1}{2^{l+1}l} \sum_{k=0}^{l} \binom{l}{k} \sum_{\mu=1}^{n-1} e^{i2\pi\mu(k-\frac{1}{2}l)/n}$$

$$= \frac{1}{2^{l+1}l} \sum_{k=0}^{l} \binom{l}{k} \left\{\sum_{\mu=0}^{n-1} e^{i2\pi\mu(k-\frac{1}{2}l)/n} - 1\right\}.$$

The inner sum has obviously these values

$$\sum_{\mu=0}^{n-1} e^{i2\pi\mu(k-\frac{1}{2}l)/n} \begin{cases} = n & \text{if } k - \tfrac{1}{2}l \text{ is divisible by } n \\ = 0 & \text{otherwise.} \end{cases}$$

Also

$$\sum_{k=0}^{l} \binom{l}{k} \cdot (-1) = -2^l.$$

Consequently

$$\alpha_l = \frac{n}{2^{l+1}l} {\sum_k}' \binom{l}{k} - \frac{1}{2l},$$

where ${\sum_k}'$ extends over those $k = 0, \cdots, l$, for which $k - \frac{1}{2}l$ is divisible by n.

Let us now determine the k occurring in the following sum (as above, $k - \frac{1}{2}l$ is divisible by n) ${\sum_k}'$. $k = \frac{1}{2}l$ is clearly one of them. All others are of the form $k = \frac{1}{2}l \pm hn$, $h = 1, 2, \cdots$. The term contributed is the same for $+$ and for $-$, since

$$\binom{l}{\frac{1}{2}l + hn} = \binom{l}{\frac{1}{2}l - hn}.$$

So we have

$$\alpha_l \begin{cases} = 0, & \text{for } l \text{ odd,} \\ = \dfrac{1}{2l}\left\{\dfrac{n}{2^l}\left[\dbinom{l}{\frac{1}{2}l} + 2 \displaystyle\sum_{h=1,2,\cdots} \dbinom{l}{\frac{1}{2}l - hn}\right] - 1\right\}, & \text{for } l \text{ even.} \end{cases}$$

[7] As pointed out above, we need to consider only the even l.

The number of terms which the sum $\sum_{h=1,2,\cdots}$ contributes depends on the comparative sizes of l and n. The number is clearly

$$\begin{array}{ll} 0 \text{ for} & \tfrac{1}{2}l < n, \\ 1 \text{ for} & n \leqq \tfrac{1}{2}l < 2n, \\ 2 \text{ for} & 2n \leqq \tfrac{1}{2}l < 3n, \\ \cdots & \cdots\cdots\cdots\cdots . \end{array}$$

Explicit formulae follow:[8]

$$\alpha_1 = \alpha_3 = \alpha_5 = \alpha_7 = \alpha_9 = \cdots = 0,$$

$$\alpha_2 = \frac{n-2}{8}, \qquad (0 \text{ for } n = 1),$$

$$\alpha_4 = \frac{3n-8}{64}, \qquad (0 \text{ for } n = 1, 2),$$

$$\alpha_6 = \frac{5n-16}{192}, \qquad \left(0 \text{ for } n = 1, 2; \frac{1}{384}, n = 3\right),$$

$$\alpha_8 = \frac{35n-128}{2048}, \quad \left(0 \text{ for } n = 1, 2; \frac{1}{2048}, n = 3; \frac{1}{128}, n = 4\right).$$

$$\beta_1 = \beta_3 = \beta_5 = \beta_7 = \beta_9 = \cdots = 0,$$

$$\beta_2 = \frac{n-2}{8}, \qquad (0 \text{ for } n = 1),$$

$$\beta_4 = \frac{n^2+2n-12}{128}, \qquad (0 \text{ for } n = 1, 2),$$

$$\beta_6 = \frac{n^3+12n^2+8n-168}{3072}, \qquad \left(0 \text{ for } n = 1, 2; \frac{5}{1024}, n = 3\right),$$

$$\beta_8 = \frac{n^4+28n^3+212n^2-64n-3696}{98304}, \quad \left(0 \text{ for } n = 1, 2; \frac{35}{32768}, n = 3; \frac{35}{2048}, n = 4\right).$$

$$M_1 = M_3 = M_5 = M_7 = M_9 = \cdots = 0,$$

$$M_2 = \frac{8}{(n-1)(n+1)} \cdot \beta_2 = \frac{n-2}{(n-1)(n+1)}, \qquad (0 \text{ for } n = 1),$$

---

[8] The author wishes to express his thanks to Miss B. I. Hart for her kind help in carrying out these computations.

$$M_4 = \frac{384}{(n-1)(n+1)(n+3)(n+5)} \cdot \beta_4 = \frac{3(n^2+2n-12)}{(n-1)(n+1)(n+3)(n+5)},$$

(0 for $n = 1, 2$),

$$M_6 = \frac{46080}{(n-1)(n+1)(n+3)(n+5)(n+7)(n+9)} \cdot \beta_6$$

$$= \frac{15(n^3+12n^2+8n-168)}{(n-1)(n+1)(n+3)(n+5)(n+7)(n+9)},$$

(0 for $n = 1, 2$; $\frac{5}{1024}$, $n = 3$).

$$M_8 = \frac{10321920}{(n-1)(n+1)(n+3)(n+5)(n+7)(n+9)(n+11)(n+13)} \cdot \beta_8$$

$$= \frac{105(n^4+28n^3+212n^2-64n-3696)}{(n-1)(n+1)(n+3)(n+5)(n+7)(n+9)(n+11)(n+13)},$$

(0 for $n = 1, 2$; $\frac{35}{32768}$, $n = 3$; $\frac{112}{218789}$, $n = 4$).

We conclude this section by obtaining asymptotic formulae for the distribution of ϵ when $n \to \infty$.

In this case our formulae show that all α_l (l even) behave asymptotically like constant multiples of n. It also appears from our formulae for the β_l (l even), that

$$\beta_l = \frac{1}{(\frac{1}{2}l)!}\alpha_2^{\frac{1}{2}l} + \text{a polynomial in } \alpha_2, \alpha_4, \cdots, \alpha_{l-2} \text{ of total order} \leqq \tfrac{1}{2}l - 1.$$

Consequently $\frac{1}{(\frac{1}{2}l)!}\alpha_2^{\frac{1}{2}l}$ is the dominant term in this expression, and so we have asymptotically

$$\beta_l \sim \frac{1}{(\frac{1}{2}l)!}\alpha_2^{\frac{1}{2}l} \sim \frac{1}{(\frac{1}{2}l)!}\left(\frac{n}{8}\right)^{\frac{1}{2}l}.$$

From this

$$M_l \sim \frac{l!}{\left(\frac{n}{2}\right)^l}\beta_l \sim \frac{l!}{(\frac{1}{2}l)!}\left(\frac{1}{2n}\right)^{\frac{1}{2}l}.$$

Now the normal distribution

$$c_1 e^{-y^2/2\sigma_1^2}\,dy, \qquad \left(c_1 = \frac{1}{\sigma_1\sqrt{2\pi}}\right),$$

with the mean 0 and the standard deviation σ_1 has the moments

$$m_l = \int_{-\infty}^{\infty} y^l c_1 e^{-y^2/2\sigma_1^2}\,dy.$$

This is clearly 0 for l odd, while for l even[9]

$$m_l = \sigma_1^{l+1} c_1 \cdot 2^{\frac{1}{2}(l+1)} \int_0^\infty e^{-u} u^{\frac{1}{2}(l-1)} du$$

$$= 2^{\frac{1}{2}(l+1)} \sigma_1^{l+1} c_1 \Gamma\left(\frac{l+1}{2}\right).$$

For $l = 0$ this becomes, since $m_0 = 1$,

$$1 = 2^{\frac{1}{2}} \sigma_1 c_1 \Gamma(\tfrac{1}{2}).$$

Dividing the former equation by the latter gives, since

$$\frac{\Gamma\left(\frac{l+1}{2}\right)}{\Gamma(\frac{1}{2})} = \frac{1}{2} \cdot \frac{3}{2} \cdots \frac{l-1}{2},$$

$$m_l = 1 \cdot 3 \cdots (l-1)\sigma_1^l = \frac{l!}{2^{\frac{1}{2}l}(\frac{1}{2}l)!}\sigma_1^l = \frac{l!}{(\frac{1}{2}l)!}\left(\frac{\sigma_1^2}{2}\right)^{\frac{1}{2}l}.$$

Comparing the formulae for M_l and for m_l shows that $M_l \sim m_l$ if $\frac{1}{2n} = \frac{\sigma_1^2}{2}$, $\sigma_1 = \sqrt{\frac{1}{n}}$. So we see:

For $n \to \infty$ the distribution of ϵ becomes asymptotically normal, with the mean 0 and the standard deviation $\sigma_1 = \sqrt{\frac{1}{n}}$. (The same result could be obtained by applying the general theorems of Liapounoff and others.)

5. The distribution law, general discussion. We return to the quantity γ, defined at the beginning of the preceding section, of which our ϵ is a special case. We wish to obtain direct information concerning the distribution law of this γ.

Since a permutation of the B_μ is permissible, we arrange them such that

$$B_1 \geqq B_2 \geqq \cdots \geqq B_m.$$

(In the special case $\gamma = \epsilon$, the $B_\mu = \cos \frac{\mu\pi}{n}$ are given in this arrangement.)

The distribution of γ covers obviously the interval

$$B_1 \geqq y \geqq B_m.$$

And if not $B_1 = \cdots = B_m$, i.e. if $B_1 > B_m$, which we assume to be the case, then we have obviously a continuous distribution law for γ in this interval. We denote it by $\omega(y)\, dy$.

---

[9] Introduce the new integration variable $u = \frac{y^2}{2\sigma_1^2}$.

Assume for the moment that $B_m > 0$. Then the quantity

$$\gamma^{-\frac{1}{2}m} = \left(\sum_{\mu=1}^{m} B_\mu x_\mu^2\right)^{-\frac{1}{2}m},$$

is bounded, and we can therefore form its mean value. This is the $-\dfrac{m}{2}$ moment of γ (cf. the beginning of the preceding section)

$$M_{-\frac{1}{2}m} = \overline{\gamma^{-\frac{1}{2}m}} = \overline{\left(\sum_{\mu=1}^{m} B_\mu x_\mu^2\right)^{-\frac{1}{2}m}}$$
$$= \int_{B_m}^{B_1} y^{-\frac{1}{2}m}\,\omega(y)\,dy.$$

With any two $a > b > 0$ $\left(\text{we shall have } \dfrac{a}{b} \to \infty \text{ subsequently}\right)$ form the quantity

$$t(a, b) = \underset{a^2 \geq \sum_{\mu=1}^{m} x_\mu^2 \geq b^2}{\int \cdots \int} \left(\sum_{\mu=1}^{m} x_\mu^2\right)^{-\frac{1}{2}m} dx_1 \cdots dx_m \text{ [10]}$$
$$= \int_b^a r^{-m}\cdot\Sigma_m\, r^{m-1}\,dr = \Sigma_m \int_b^a \frac{dr}{r}$$
$$= \Sigma_m \ln \frac{a}{b}.$$

Consider next

$$s(a, b) = \underset{a^2 \geq \sum_{\mu=1}^{m} \frac{1}{B_\mu} x_\mu^2 \geq b^2}{\int \cdots \int} \left(\sum_{\mu=1}^{m} x_\mu^2\right)^{-\frac{1}{2}m} dx_1 \cdots dx_m \text{ [11]}$$
$$= \underset{a^2 \geq \sum_{\mu=1}^{m} x_\mu^2 \geq b^2}{\int \cdots \int} \left(\sum_{\mu=1}^{m} B_\mu x_\mu^2\right)^{-\frac{1}{2}m} \sqrt{\prod_{\mu=1}^{m} B_\mu}\, dx_1 \cdots dx_m$$
$$= \underset{a^2 \geq \sum_{\mu=1}^{m} x_\mu^2 \geq b^2}{\int \cdots \int} M_{-\frac{1}{2}m} \left(\sum_{\mu=1}^{m} x_\mu^2\right)^{-\frac{1}{2}m} \sqrt{\prod_{\mu=1}^{m} B_\mu}\, dx_1 \cdots dx_m$$
$$= M_{-\frac{1}{2}m} \sqrt{\prod_{\mu=1}^{m} B_\mu}\; t(a, b).$$

---

[10] Concerning this transformation to polar coordinates and the quantity Σ_m cf. the first part of the preceding section.

[11] Replace each variable x_μ by $\sqrt{B_\mu}\, x_\mu$.

On the other hand, a comparison of their respective integration domains makes it clear that

$$t(B_m a, B_1 b) \leqq s(a, b) \leqq t(B_1 a, B_m b).$$

Thus

$$\Sigma_m \ln \frac{B_m a}{B_1 b} \leqq M_{-\frac{1}{2}m} \sqrt{\prod_{\mu=1}^{m} B_\mu} \cdot \Sigma_m \ln \frac{a}{b} \leqq \Sigma_m \ln \frac{B_1 a}{B_m b},$$

i.e.

$$\frac{1}{\sqrt{\prod_{\mu=1}^{m} B_\mu}} \frac{\ln \frac{a}{b} - \ln \frac{B_1}{B_m}}{\ln \frac{a}{b}} \leqq M_{-\frac{1}{2}m} \leqq \frac{1}{\sqrt{\prod_{\mu=1}^{m} B_\mu}} \frac{\ln \frac{a}{b} + \ln \frac{B_1}{B_m}}{\ln \frac{a}{b}}.$$

Now let $\frac{a}{b} \rightarrow \infty$, then

$$M_{-\frac{1}{2}m} = \frac{1}{\sqrt{\prod_{\mu=1}^{m} B_\mu}}$$

obtains, i.e.

$$\overline{\gamma^{-\frac{1}{2}m}} = \int_{B_m}^{B_1} y^{-\frac{1}{2}m} \omega(y)\, dy = \frac{1}{\sqrt{\prod_{\mu=1}^{m} B_\mu}}.$$

We now drop the assumption $B_m > 0$. We consider instead a real number z with $z < B_m$. Replace each B_μ by $B_\mu - z$. Then the one with $\mu = m$ becomes > 0. And γ is obviously replaced by $\gamma - z$. Consequently our above equation is now valid in the form

$$\overline{(\gamma - z)^{-\frac{1}{2}m}} = \int_{B_m}^{B_1} (y - z)^{-\frac{1}{2}m} \omega(y)\, dy = \frac{1}{\sqrt{\prod_{\mu=1}^{m} (B_\mu - z)}}.$$

Let now z be a complex variable. The second term of the above equation is a (locally) analytical function of z, except in the (real) interval $B_1 \geqq z \geqq B_m$. The third term, too, is a (locally) analytical function of z, except at the (real) points $B_1, \cdots, B_m$. Consequently both are one-valued analytical functions of z in the simply connected domain which obtains from the complex z plane by exclusion of the (real) half line

$$z \geqq B_m.$$

Hence the equation

$$\text{(1)} \qquad \int_{B_m}^{B_1} (y - z)^{-\frac{1}{2}m} \omega(y)\, dy = \frac{1}{\sqrt{\prod_{\mu=1}^{m} (B_\mu - z)}},$$

which holds for all (real) $z < B_m$, remains true for all complex z of the above domain.[12]

We observe next that $\omega(y)$ is an analytical function of y in $B_1 \geqq y \geqq B_m$, whenever $y \neq B_1, \cdots, B_m$. This is easily established by using any multiple integral expression for $\omega(y)$ which, while hard to evaluate explicitly, puts this analyticity into evidence.[13]

---

[12] $(y - z)^{-\frac{1}{2}m}$ and the factors $(B_\mu - z)^{-\frac{1}{2}}$ of $\dfrac{1}{\sqrt{\prod_{\mu=1}^{m} (B_\mu - z)}}$ are those branches of these analytical functions which are (real and) > 0 when z is (real and) $< B_\mu$. When m is even (as it will be, cf. below) the domain of analyticity is somewhat more extended, but we need not discuss this.

[13] The computation which follows gives the desired analyticity in a simple way, and also makes it clear why the analyticity fails at $y = B_1, \cdots, B_m$.

Consider the $y \neq B_1, \cdots, B_m$ in $B_1 \geqq y \geqq B_m$. The probability of $\gamma \leqq y$ is $p(y) = \int_{B_m}^{y} \omega(y)\, dy$, and we may establish its analyticity instead of that of $p'(y) = \omega(y)$.

Obviously $p(y)$ is equally the probability of $\sum_{\mu=1}^{m} B_\mu x_\mu^2 \leqq y \sum_{\mu=1}^{m} x_\mu^2$, if the $x_1, \cdots, x_m$ are equidistributed over a spherical surface $\sum_{\mu=1}^{m} x_\mu^2 = r^2$, with any given $r > 0$.

Our hypotheses concerning y imply $B_v > y > B_{v+1}$ for a suitable $v = 1, \cdots, m - 1$.

Consider now the expression

$$f(y) = \underset{\sum_{\mu=1}^{m} B_\mu x_\mu^2 \leqq y \sum_{\mu=1}^{m} x_\mu^2}{\int \cdots \int} e^{-\sum_{\mu=1}^{m} x_\mu^2}\, dx_1 \cdots dx_m .$$

Transforming to polar coordinates, we obtain

$$\begin{aligned} f(y) &= \int_0^\infty e^{-r^2} \cdot \Sigma_m\, p(y) r^{m-1}\, dr \\ &= \Sigma_m \int_0^\infty e^{-r^2} r^{m-1}\, dr \cdot p(y). \end{aligned}$$

(Σ_m as before.) Hence it suffices to establish the analyticity of $f(y)$. Now on the other hand

$$\begin{aligned} f(y) &= \underset{\sum_{\mu=1}^{v} (B_\mu - y) x_\mu^2 \leqq \sum_{\mu=v+1}^{m} (y - B_\mu) x_\mu^2}{\int \cdots \int} e^{-\sum_{\mu=1}^{m} x_\mu^2}\, dx_1 \cdots dx_m \\ &= \frac{1}{\sqrt{\prod_{\mu=1}^{m} \vert B_\mu - y \vert}} \underset{\sum_{\mu=1}^{v} w_\mu^2 \leqq \sum_{\mu=v+1}^{m} w_\mu^2}{\int \cdots \int} e^{-\sum_{\mu=1}^{n} w_\mu^2 / \vert B_\mu - y \vert}\, dw_1 \cdots dw_m . \end{aligned}$$

(We introduced the new variables $w_\mu = \sqrt{\vert B_\mu - y \vert}\, x_\mu$.) And this expression is clearly analytical in y, since $B_v > y > B_{v+1}$.

We shall need only the fact that $\omega(y)$ possesses $\frac{1}{2}m$ continuous derivatives at these places. (m will be assumed to be even, cf. below.) Its behavior at $y = B_1, \cdots, B_m$ will follow from our subsequent results in all cases where we need it.

In order to determine $\omega(y)$ from (1), as we now propose to do, it is very convenient to assume that m is even. We therefore make this assumption, and shall maintain it throughout most of what follows.

Consider a $y_0 \neq B_1, \cdots, B_m$ in $B_1 \geqq y_0 \geqq B_m$. Then $B_v > y > B_{v+1}$ for a suitable $v = 1, \cdots, m-1$. Now put

$$z = y_0 + it \qquad (t \text{ real and } > 0),$$

form (1), take the imaginary parts of both sides, and let $t \to 0$.

Consider first the left-hand side of (1). Since $\omega(y)$ possesses $\frac{1}{2}m$ continuous derivatives at $y = y_0$, we have

$$\omega(y) = \sum_{k=0}^{\frac{1}{2}m-1} \theta_k (y - y_0)^k + e(y)(y - y_0)^{\frac{1}{2}m}$$

with a bounded $e(y)$. Clearly

$$\theta_k = \frac{1}{k!}\left\{\frac{d^k}{dy^k}\,\omega(y)\right\}_{y=y_0}.$$

Thus, since $\omega(y)$ is real, all θ_k are real and $e(y)$ is also real.

Compute now the contribution of each one of the $\frac{1}{2}m + 1$ terms in the above expression for $\omega(y)$ to the imaginary part of the left-hand side of (1).

The last term, $e(y)\cdot(y - y_0)^{\frac{1}{2}m}$, gives

$$\Im \int_{B_m}^{B_1} (y - y_0 - it)^{-\frac{1}{2}m} e(y)(y - y_0)^{\frac{1}{2}m}\,dy = \Im \int_{B_m}^{B_1} \left(\frac{y - y_0}{y - y_0 - it}\right)^{\frac{1}{2}m} e(y)\,dy.$$

The integrand is uniformly bounded, and so the reality conditions cause the entire expression to $\to 0$ for $t \to 0$. Hence the contribution of this term is zero for $t \to 0$.

The other $\dfrac{m}{2}$ terms correspond to $k = 0, 1, \cdots, \dfrac{m}{2} - 1$, the k term being

$$\begin{aligned}
\Im \int_{B_m}^{B_1} & (y - y_0 - it)^{-\frac{1}{2}m}\cdot\theta_k (y - y_0)^k \cdot dy \\
&= \theta_k \Im \int_{B_m}^{B_1} \frac{(y - y_0)^k}{(y - y_0 - it)^{\frac{1}{2}m}}\,dy \\
&= \theta_k \Im \int_{B_m}^{B_1} \frac{\sum_{h=0}^{k} \binom{k}{h} (it)^h (y - y_0 - it)^{k-h}}{(y - y_0 - it)^{\frac{1}{2}m}}\,dy \\
&= \theta_k \sum_{h=0}^{k} \binom{k}{h} \Im \left\{ (it)^h \int_{B_m}^{B_1} (y - y_0 - it)^{k-h-\frac{1}{2}m}\,dy \right\}.
\end{aligned}$$

The exponent $k - h - \frac{m}{2}$ in the integral is always $\leqq \left(\frac{m}{2} - 1\right) - 0 - \frac{m}{2} = -1$, and it is $= -1$ if and only if $k = \frac{m}{2} - 1$, $h = 0$. Consider first a term where this is not the case, i.e. where the exponent $k - h - \frac{m}{2} < -1$. For such a term the expression $\mathfrak{J}\{\cdots\}$ becomes

$$\mathfrak{J}(it)^h \frac{1}{k - h - \frac{m}{2} + 1} \{(y - y_0 - it)^{k-h-\frac{1}{2}m+1}\}_{y=B_m}^{y=B_1}.$$

For $t \to 0$ the last factors are bounded and real, and so the entire expression $\to 0$: for $h = 0$ because of the reality conditions, for $h > 0$ because of $(it)^h \to 0$. Thus only the term $k = \frac{m}{2} - 1$, $h = 0$ can contribute something else than zero for $t \to 0$.

Now this term is equal to

$$\theta_{\frac{1}{2}m-1}\mathfrak{J}\{\ln(y - y_0 - it)\}_{y=B_m}^{y=B_1},$$

and for $t \to 0$ this converges to

$$\theta_{\frac{1}{2}m-1}\mathfrak{J}(i\pi) = \pi\theta_{\frac{1}{2}m-1} = \frac{\pi}{\left(\frac{m}{2} - 1\right)!}\left\{\frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}}\omega(y)\right\}_{y=y_0}.$$ [14]

Thus the imaginary part of the entire left-hand side of (1) converges for $t \to 0$ to this expression.

The right-hand side of (1) is easier to discuss. The imaginary part under consideration is now

$$\mathfrak{J}\frac{1}{\sqrt{\prod_{\mu=1}^{m}(B_\mu - y_0 - it)}} = \mathfrak{J}\prod_{\mu=1}^{m}(B_\mu - y_0 - it)^{-\frac{1}{2}}.$$

Considering[12] (its y is our $y_0 + it$), this converges for $t \to 0$ to

$$\mathfrak{J}\prod_{\mu=1}^{v}(B_\mu - y_0)^{-\frac{1}{2}}\prod_{\mu=v+1}^{m} i(y_0 - B_\mu)^{-\frac{1}{2}} = \mathfrak{J}i^{m-v}\frac{1}{\sqrt{\prod_{\mu=1}^{m}|B_\mu - y_0|}}.$$ [15]

---

[14] This evaluation $\{\ln(y - y_0 - it)\}_{y=B_m}^{y=B_1} \to i\pi$ is based on $t > 0$, and the fact that y moves on the real axis from B_m to B_1. It has no connection with[12].

[15] The square roots of the (real and) > 0 quantities

$$B_\mu - y_0\ (\mu = 1, \cdots, v), \qquad y_0 - B_\mu\ (\mu = v + 1, \cdots, m), \quad \text{and} \quad \prod_{\mu=1}^{m}|B_\mu - y_0|$$

are taken to be > 0.

If v (hence $m - v$) is even, then this is zero. If v (hence $m - v$) is odd, then this is equal to $(-1)^{\frac{1}{2}(m-v-1)} \dfrac{1}{\sqrt{\prod_{\mu=1}^{m} | B_\mu - y_0 |}}$. Thus (1) becomes the following equation:

$$\frac{\pi}{\left(\frac{m}{2} - 1\right)!} \left\{ \frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}} \omega(y) \right\}_{y=y_0} \begin{cases} = 0 & \text{if } v \text{ is even,} \\ = (-1)^{\frac{1}{2}(m-v-1)} \dfrac{1}{\sqrt{\prod_{\mu=1}^{m} | B_\mu - y_0 |}} & \text{if } v \text{ is odd.} \end{cases}$$

Simplifying this, and writing y for y_0, and also restating the definition of v gives

$$\frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}} \omega(y) \begin{cases} = 0 & \text{if } v \text{ is even,} \\ = (-1)^{\frac{1}{2}(m-v-1)} \dfrac{\left(\frac{m}{2} - 1\right)!}{\pi} \dfrac{1}{\sqrt{\prod_{\mu=1}^{m} | B_\mu - y |}} & \text{if } v \text{ is odd,} \end{cases} \tag{2}$$

$$B_v > y > B_{v+1},\ v = 1, \cdots, m - 1.$$

Observe finally, that if we put

$$\mathfrak{A}(y) = \prod_{\mu=1}^{m} (y - B_\mu),$$

then this product has v factors < 0 ($\mu = 1, \cdots, v$), while the others are > 0. So

$$\mathfrak{A}(y) \gtrless 0 \qquad \text{for} \qquad v \ \begin{matrix}\text{even}\\ \text{odd}\end{matrix},$$

and in the latter case

$$\prod_{\mu=1}^{m} | B_\mu - y | = -\mathfrak{A}(y).$$

It is clear how we may now rewrite (2).

We are now in a position to determine the behavior of $\omega(y)$ at $y = B_1, \cdots, B_m$ too, since we know how its $\frac{m}{2} - 1$-th derivative behaves in the immediate vicinity of these places. (2) shows that it is singular there, and that the nature of the singularity depends on the number of the μ, for which B_μ is equal to the y in question, i.e. on the multiplicity of this root of our polynomial $\mathfrak{A}(y)$.

In our actual application (to $\gamma = \epsilon$, cf. the beginning of this section) the

B_μ are pairwise different, i.e. all root multiplicities of $\mathfrak{A}(y)$ are equal to one. A further special case, which has a certain interest of its own, is when the B_μ are equal two by two, but otherwise different, i.e. all root multiplicities of $\mathfrak{A}(y)$ are equal to two. In the discussion which follows we shall therefore assume that one or the other of these two cases occurs.

In the first case $\frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}}\,\omega(y)$ has on each side of a $y = B_\mu$ one of these two behaviors: It is identically zero, or it is singular, of the type $\frac{1}{\sqrt{|B_\mu - y|}}$. Thus it is at any rate integrable. Consequently $\frac{d^{\frac{1}{2}m-2}}{dy^{\frac{1}{2}m-2}}\,\omega(y)$ is continuous on each side of $y = B_\mu$, i.e. for both $y = B_\mu \pm 0$. Successive integrations give now that all $\frac{d^k}{dy^k}\,\omega(y)$, $k = 0, 1, \cdots, \frac{m}{2} - 2$, are continuous for both $y = B_\mu \pm 0$.

In the second case we have $B_1 = B_2 > B_3 = B_4 > \cdots > B_{m-1} = B_m$. So the v with $B_v > y > B_{v+1}$ is necessarily even, and $\frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}}\,\omega(y)$ is identically zero for all of (2). Consequently $\frac{d^{\frac{1}{2}m-2}}{dy^{\frac{1}{2}m-2}}\,\omega(y)$ is again continuous on each side of $y = B_\mu$, i.e. for both $y = B_\mu \pm 0$. Successive integrations show again that all $\frac{d^k}{dy^k}\,\omega(y)$, $k = 0, 1, \cdots, \frac{m}{2} - 2$, are continuous for both $y = B_\mu \pm 0$.

We must therefore discuss only how much the $\frac{d^k}{dy^k}\,\omega(y)$, $k = 0, 1, \cdots, \frac{m}{2} - 2$, change from $y = B_\mu - 0$ to $y = B_\mu + 0$.

Let us return to the procedure by which we derived (2) from (1). We put again

$$z = y_0 + it \qquad (t \text{ real and } > 0)$$

and let $t \to \infty$. But we consider now (1) itself (and not merely its imaginary part), and we choose a $y_0 = B_v$.

Consider first the left-hand side of (1), always disregarding terms which stay bounded for $t \to 0$. Then we can replace the integral $\int_{B_m}^{B_1}$ of (1) by any $\int_{B_v-a}^{B_v+a}$ with any fixed $a > 0$, and this is equal to

$$\int_{B_v-a}^{B_v-0} + \int_{B_v+0}^{B_v+a}.$$

We choose this $a > 0$ so small that no $B_\mu \neq B_v$ lies between $B_v - a$ and $B_v + a$. I.e. all $\frac{d^k}{dy^k}\,\omega(y)$, $k = 0, 1, \cdots, \frac{m}{2} - 2$, are continuous from $B_v - a$ to $B_v - 0$ and also from $B_v + 0$ to $B_v + a$.

This being the case, we can evaluate the above sum of two integrals by $\frac{m}{2} - 1$ successive partial integrations. Thus we get

$$-\left\{\sum_{k=0}^{\frac{1}{2}m-2} \frac{\left(\frac{m}{2} - 2 - k\right)!}{\left(\frac{m}{2} - 1\right)!} (y - B_v - it)^{-\frac{1}{2}m+1+k} \frac{d^k}{dy^k}\,\omega(y)\right\}_{y=B_v-a}^{y=B_v-0}$$

$$-\left\{\sum_{k=0}^{\frac{1}{2}m-2} \frac{\left(\frac{m}{2} - 2 - k\right)!}{\left(\frac{m}{2} - 1\right)!} (y - B_v - it)^{-\frac{1}{2}m+1+k} \frac{d^k}{dy^k}\,\omega(y)\right\}_{y=B_v+0}^{y=B_v+a}$$

$$+ \frac{1}{\left(\frac{m}{2} - 1\right)!} \int_{B_v-a}^{B_v+a} (y - B_v - it)^{-1} \frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}}\,\omega(y)\,dy.$$

In the first two lines the $y = B_v \pm a$ terms are bounded for $t \to 0$, therefore only the $y = B_v \pm 0$ terms need be kept. Then the first two lines give

$$\sum_{k=0}^{\frac{1}{2}m-2} \frac{\left(\frac{m}{2} - 2 - k\right)!}{\left(\frac{m}{2} - 1\right)!} (-it)^{-\frac{1}{2}m+1+k} \left\{\frac{d^k}{dy^k}\,\omega(y)\right\}_{y=B_v-0}^{y=B_v+0},$$

up to terms which stay bounded for $t \to 0$. Consider now the third line. We know that the $\frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}}\,\omega(y)$ in its integrand can be majorized by $\frac{c_2}{\sqrt{|y - B_v|}}$ (for a suitable constant c_2, cf. our discussion preceding the present one). Thus the integral in question is majorized by

$$\int_{B_v-a}^{B_v+a} |y - B_v - it|^{-1} c_2 |y - B_v|^{-\frac{1}{2}}\,dy,$$

hence *a fortiori* by

$$\int_{-\infty}^{\infty} |y - B_v - it|^{-1} c_2 |y - B_v|^{-\frac{1}{2}}\,dy \text{ [16]}$$

$$= c_2 t^{-\frac{1}{2}} \int_{-\infty}^{\infty} |u - i|^{-1} |u|^{-\frac{1}{2}}\,du$$

$$= c_2 t^{-\frac{1}{2}} \int_{-\infty}^{\infty} \frac{du}{\sqrt{(u^2 + 1)\cdot|u|}} \text{ [17]}$$

$$= c_2 \int_0^{\infty} \frac{dv}{\sqrt{v^4 + 1}}\, t^{-\frac{1}{2}}.$$

---

[16] Introduce the new integration variable $u = \dfrac{y - B_v}{t}$.

[17] Introduce the new integration variable $v = \sqrt{|u|}$.

Since the last integration is obviously finite, the entire expression is $O(t^{-\frac{1}{2}})$ for $t \to 0$.

Consequently the left-hand side of (1) is equal to

$$\sum_{k=0}^{\frac{1}{2}m-2} \frac{\left(\frac{m}{2} - 2 - k\right)!}{\left(\frac{m}{2} - 1\right)!} (-it)^{-\frac{1}{2}m+1+k} \left\{\frac{d^k}{dy^k} \omega(y)\right\}_{y=B_v-0}^{y=B_v+0} + O(t^{-\frac{1}{2}}),$$

for $t \to 0$. $\left(\text{For } B_v = B_1 \text{ or } B_m \text{ the } \frac{d^k}{dy^k}\omega(y) \text{ at } y = B_v + 0 \text{ or } B_v - 0\text{, respectively, must obviously be taken to be zero.}\right)$

Consider now the right-hand side of (1).

We first suppose the B_μ are pairwise different. The right-hand side in question is $\dfrac{1}{\sqrt{\prod_{\mu=1}^{m} (B_\mu - B_v - it)}}$, i.e. $O(t^{-\frac{1}{2}})$.

Secondly let us consider $B_1 = B_2 > B_3 = B_4 > \cdots > B_{m-1} = B_m$. So we may assume $v = 2\lambda$ $\left(\lambda = 1, \cdots, \frac{m}{2}\right)$. The right-hand side of (1) becomes now a rational function, $\dfrac{1}{\prod_{k=1}^{\frac{1}{2}m} (B_{2k} - z)}$. (The sign is determined by[12].) So in our case it is $\dfrac{1}{\prod_{k=1}^{\frac{1}{2}m} (B_{2k} - B_{2\lambda} - it)}$, i.e. $\dfrac{(-1)^{\frac{1}{2}m-\lambda}}{\prod_{k=1}^{\lambda-1} (B_{2k} - B_{2\lambda}) \cdot \prod_{k=\lambda+1}^{\frac{1}{2}m} (B_{2\lambda} - B_{2k})} (-it)^{-1} + O(1)$.

Comparing these with our above expression gives therefore (for $t \to 0$)

$$\sum_{k=0}^{\frac{1}{2}m-2} \frac{\left(\frac{m}{2} - 2 - k\right)!}{\left(\frac{m}{2} - 1\right)!} (-it)^{-\frac{1}{2}m+1+k} \left\{\frac{d^k}{dy^k} \omega(y)\right\}_{y=B_v-0}^{y=B_v+0}$$

$$= O(t^{-\frac{1}{2}}) \text{ in the first case,}$$

$$= \frac{(-1)^{\frac{1}{2}m-\lambda}}{\prod_{k=1}^{\lambda-1} (B_{2k} - B_{2\lambda}) \cdot \prod_{k=\lambda+1}^{\frac{1}{2}m} (B_{2\lambda} - B_{2k})} (-it)^{-1} + O(t^{-\frac{1}{2}}) \text{ in the second case.}$$

In this formula the left-hand side is a polynomial in $(-it)^{-1}$. Hence the $O(t^{-\frac{1}{2}})$ terms on the right-hand side must vanish, and otherwise all powers of $-it$ must have the same coefficient on both sides. Consequently

$$\frac{\left(\frac{m}{2} - 2 - k\right)!}{\left(\frac{m}{2} - 1\right)!} \left\{\frac{d^k}{dy^k} \omega(y)\right\}_{y=B_v-0}^{y=B_v+0}$$

must vanish, except in the second case for the one value of k with $-\frac{m}{2} + 1 + k = -1$, i.e. $k = \frac{m}{2} - 2$. So, with this one exception, we have

$$\left\{\frac{d^k}{dy^k}\,\omega(y)\right\}_{y=B_v+0} = \left\{\frac{d^k}{dy^k}\,\omega(y)\right\}_{y=B_v-0}.$$

And in the exceptional case (second case, $v = 2\lambda$)

$$\left\{\frac{d^{\frac{1}{2}m-2}}{dy^{\frac{1}{2}m-2}}\,\omega(y)\right\}_{y=B_v-0}^{y=B_v+0} = (-1)^{\frac{1}{2}m-\lambda}\left(\frac{m}{2}-1\right)!\,\frac{1}{\prod_{k=1}^{\lambda-1}(B_{2k}-B_{2\lambda})\prod_{k=\lambda+1}^{\frac{1}{2}m}(B_{2\lambda}-B_{2k})}.$$

Thus in the first case all derivatives $\frac{d^k}{dy^k}\,\omega(y)$, $k = 0, 1, \cdots, \frac{m}{2} - 2$, are continuous even at $y = B_1, \cdots, B_m$.

In the second case the same is true for $k = 0, 1, \cdots, \frac{m}{2} - 3$, but the derivative with $k = \frac{m}{2} - 2$ behaves differently for $y = B_1, \cdots, B_m$. Indeed, for $y = B_{2\lambda-1} = B_{2\lambda}$ $\left(\lambda = 1, \cdots, \frac{m}{2}\right)$ this derivative is continuous for both $y = B_{2\lambda} \pm 0$, but it increases from $B_{2\lambda} - 0$ to $B_{2\lambda} + 0$ by

$$(-1)^{\frac{1}{2}m-\lambda}\left(\frac{m}{2}-1\right)!\,\frac{1}{\prod_{k=1}^{\lambda-1}(B_{2k}-B_{2\lambda})\prod_{k=\lambda+1}^{\frac{1}{2}m}(B_{2\lambda}-B_{2k})}.$$

(At $y = B_1 + 0$ and $B_m - 0$ the $\frac{d^k}{dy^k}\,\omega(y)$ must be thought to continue with the value zero.)

These rules, together with (2), determine $\omega(y)$ completely.

6. First special case. We consider the first special case, where the B_μ are pairwise different. We immediately specialize further, to $\gamma = \epsilon$, i.e. $m = n - 1$, $B_\mu = \cos\frac{\mu\pi}{n}$ $(\mu = 1, \cdots, n-1)$. (Cf. the beginning of the preceding section.) Since m must be even, n must be odd. The rules of section 5 determine $\omega(y)$; in particular all derivatives $\frac{d^k}{dy^k}\,\omega(y)$, $k = 0, 1, \cdots, \frac{n-1}{2} - 2$, are everywhere continuous, beginning and ending with zero at $y = B_1$ and B_{n-1}, respectively.

In the even intervals

$$B_2 \geqq y \geqq B_3, \qquad B_4 \geqq y \geqq B_5, \cdots, B_{n-3} \geqq y \geqq B_{n-2},$$

the derivative $\frac{d^{\frac{1}{2}(n-1)-1}}{dy^{\frac{1}{2}(n-1)-1}}\,\omega(y)$ is zero, i.e. $\omega(y)$ is a polynomial of degree $\frac{1}{2}(n-1)-2$. In the odd intervals

$$B_1 \geqq y \geqq B_2, \qquad B_3 \geqq y \geqq B_4, \cdots, B_{n-2} \geqq y \geqq B_{n-1},$$

we have

$$\frac{d^{\frac{1}{2}(n-1)-1}}{dy^{\frac{1}{2}(n-1)-1}}\,\omega(y) = \pm\frac{(\frac{1}{2}[n-1]-1)!}{\pi}\,\frac{1}{\sqrt{-\mathfrak{A}(y)}}$$

(the sign $\pm$ is alternating $(-1)^{\frac{1}{2}(n-1)-1}, (-1)^{\frac{1}{2}(n-1)-2}, \cdots, +$), where

$$\mathfrak{A}(y) = \prod_{\mu=1}^{n-1}\left(y - \cos\frac{\mu\pi}{n}\right).$$

Another expression for $\mathfrak{A}(y)$ may be found by the following method.

Clearly $$\frac{\sin(n\varphi)}{\sin\varphi} = \frac{e^{in\varphi}-e^{-in\varphi}}{e^{i\varphi}-e^{-i\varphi}} = \sum_{\mu=0}^{n-1} e^{i(n-1-2\mu)\varphi}$$

is a polynomial of $\cos\varphi = \frac{1}{2}(e^{i\varphi}+e^{-i\varphi})$ of degree $n-1$, with the highest coefficient 2^{n-1}. For $\varphi = \frac{\mu\pi}{n}$, $\mu = 1, \cdots, n-1$, $\sin(n\varphi) = 0$, $\sin\varphi \neq 0$, hence $\frac{\sin(n\varphi)}{\sin\varphi}$, as a polynomial in $\cos\varphi$, has the same roots as $\mathfrak{A}(y)$. $\mathfrak{A}(y)$ is a polynomial of degree $n-1$ with the highest coefficient 1. Consequently

$$\mathfrak{A}(\cos\varphi) = \frac{1}{2^{n-1}}\,\frac{\sin(n\varphi)}{\sin\varphi}.$$

This formula allows one to compute $\mathfrak{A}(y)$ quickly, examples are

$$n = 3\colon\ \mathfrak{A}(y) = y^2 - \tfrac{1}{4},$$
$$n = 5\colon\ \mathfrak{A}(y) = y^4 - \tfrac{3}{4}y^2 + \tfrac{1}{16},$$
$$n = 7\colon\ \mathfrak{A}(y) = y^6 - \tfrac{5}{4}y^4 + \tfrac{3}{8}y^2 - \tfrac{1}{64}.$$

The number of odd intervals, on which integrations must be carried out, is $\frac{1}{2}(n-1)$, but since those which are symmetric with respect to 0 require the same computations, only $\frac{1}{4}(n-1)$ or $\frac{1}{4}(n+1)$ must be considered. So there are $1, 1, 2, \cdots$ such intervals for $n = 3, 5, 7, \cdots$ respectively. The integrals are first elementary (arcsin), then elliptic, then hyperelliptic.

Numerical computations for $n = 3$ are immediate; for $n = 5, 7$ they have been carried out with considerable precision by B. I. Hart.

At $y = B_\mu$, $\frac{d^{\frac{1}{2}(n-1)-1}}{dy^{\frac{1}{2}(n-1)-1}}\,\omega(y)$ has a singularity of the type $\frac{1}{\sqrt{|y-B_\mu|}}$ (cf. the end of section 5), while all $\frac{d^k}{dy^k}\,\omega(y)$, $k = 0, 1, \cdots, \frac{1}{2}(n-1)-2$, are continuous.

At $y = B_1$ and B_{n-1}, in particular, they are zero. Hence it follows by successive integrations that the order of vanishing of $\frac{d^k}{dy^k}\,\omega(y)$, $k = 0, 1, \cdots, \frac{1}{2}(n-1) - 2$ at $y = B_1$ and B_{n-1} is $(\frac{1}{2}(n-1) - 1) - k - \frac{1}{2} = \frac{n}{2} - 2 - k$. In particular for $k = 0$ we find that at its maximum and at its minimum $\left(B_1 \text{ and } B_{n-1}\text{, i.e. } \pm \cos\frac{\pi}{n}\right)$ the order of vanishing of $\omega(y)$ is $\frac{n}{2} - 2$.[18]

Since $\omega(y)$ has this property, and since it is obviously an even function of y, R. H. Kent has suggested approximating it by a series expansion of the form

$$\omega(y) = \sum_{h=0}^{\infty} a_h \left(\cos^2\frac{\pi}{n} - y^2\right)^{\frac{1}{2}n-2+h}. \tag{3}$$

Computations by B. I. Hart, not yet published, have shown that even the use of the first four terms ($h = 0, 1, 2, 3$, the a_h being determined by the condition of normalization and by the first three even moments of the actual distribution given in section 4) give excellent approximations. The use of the formula (3) suggests itself likewise for even values of n.

7. Second special case. We consider now the second special case, where $B_1 = B_2 > B_3 = B_4 > \cdots > B_{m-1} = B_m$. This has no immediate bearing on our original problem (cf. the preceding section), but we shall nevertheless discuss it for the two following reasons. First, it is hoped that the reader will find an independent interest in the simple and complete results which can be obtained in this case. Second, there are various modifications of our original problem, which lead to this case. For example let the $x_1, \cdots, x_n$ in our original problem, as described in section 1, be complex numbers instead of real ones, replacing all squares by absolute value squares. Then one verifies easily that all characteristic values $\lambda_1, \cdots, \lambda_{n-1}$ are doubled, and so our first case goes over into our second case. (This amounts to replacing our quadratic forms by Hermitian forms, cf.[4]) It is easy to imagine two-dimensional problems where this set-up is natural.

We put $C_\lambda = B_{2\lambda-1} = B_{2\lambda}$ for $\lambda = 1, \cdots, \frac{m}{2}$, so that $C_1 > C_2 > \cdots > C_{\frac{1}{2}m}$ are the only restrictions imposed.

Every y in $B_1 \geqq y \geqq B_m$, i.e. in $C_1 \geqq y \geqq C_{\frac{1}{2}m}$, lies in an interval $C_\lambda \geqq y \geqq C_{\lambda+1}$ i.e. $B_{2\lambda} \geqq y \geqq B_{2\lambda+1}$. That is the v of (2) is always even, and so $\frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}}\,\omega(y)$ is zero in every one of these intervals. Therefore $\omega(y)$ is a polynomial of degree $\frac{m}{2} - 2$ in every one of these intervals. We have already shown that $\omega(y)$ is

[18] We omit the simple discussion of $n = 3$, which must be excluded from this result.

not the same polynomial in each interval. Thus $\omega(y)$ is represented by $\frac{m}{2} - 1$ polynomials of degree $\frac{m}{2} - 2$ in the $\frac{m}{2} - 1$ intervals

$$C_1 \geqq y \geqq C_2, \quad C_2 \geqq y \geqq C_3, \cdots, C_{\frac{1}{2}m-1} \geqq y \geqq C_{\frac{1}{2}m}.$$

We could try to obtain explicit expressions for these polynomials by a direct application of the results at the close of section 5. A characterization of the distribution can, however, be obtained in a more elegant way by an indirect procedure.

Consider an arbitrary function $\mathfrak{F}(y)$. We wish to express its mean

$$\overline{\mathfrak{F}(y)} = \int_{C_{\frac{1}{2}m}}^{C_1} \mathfrak{F}(y)\omega(y)\,dy.$$

If we can do this for all $\mathfrak{F}(y)$ then the distribution is completely characterized.

We select first an $\frac{m}{2} - 1$-fold primitive function of $\mathfrak{F}(y)$, i.e. a function $\mathfrak{G}(y)$ with

$$\frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}}\,\mathfrak{G}(y) = \mathfrak{F}(y).$$

Of course $\mathfrak{G}(y)$ is determined only up to an additive polynomial of degree $\frac{m}{2} - 2$ in y.

Now the above expectation value becomes

$$\begin{aligned}\overline{\mathfrak{F}(y)} &= \int_{C_{\frac{1}{2}m}}^{C_1} \frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}}\,\mathfrak{G}(y)\omega(y)\,dy \\ &= \sum_{\lambda=1}^{\frac{1}{2}m-1} \int_{C_{\lambda+1}+0}^{C_\lambda-0} \frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}}\,\mathfrak{G}(y)\omega(y)\,dy.\end{aligned}$$

Since all $\frac{d^k}{dy^k}\,\omega(y)$, $k = 0, 1, \cdots, \frac{m}{2} - 2$, are continuous from $C_{\lambda+1} + 0$ to $C_\lambda - 0$ for all $\lambda = 1, \cdots, \frac{m}{2} - 1$, we can evaluate each integral of the above sum by $\frac{m}{2} - 1$ successive partial integrations. Thus the following expression obtains:

$$\sum_{\lambda=1}^{\frac{1}{2}m-1} \left\{ \sum_{k=0}^{\frac{1}{2}m-2} (-1)^k \frac{d^{\frac{1}{2}m-k-2}}{dy^{\frac{1}{2}m-k-2}}\,\mathfrak{G}(y)\,\frac{d^k}{dy^k}\,\omega(y) \right\}_{y=C_{\lambda+1}+0}^{y=C_\lambda-0} + (-1)^{\frac{1}{2}m-1} \int_{C_{\frac{1}{2}m}}^{C_1} \mathfrak{G}(y)\,\frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}}\,\omega(y)\,dy.$$

Considering the definition of $\mathfrak{G}(y)$ as an $\frac{m}{2} - 1$-fold primitive function, the $\frac{d^{k'}}{dy^{k'}}\,\mathfrak{G}(y)$, $k' = 0, 1, \cdots, \frac{m}{2} - 2$, are everywhere continuous. This corresponds

to $k' = \frac{m}{2} - k - 2, k = 0, 1, \cdots, \frac{m}{2} - 2$. Hence the first line can be rewritten as

$$-\sum_{\lambda=1}^{\frac{1}{2}m} \sum_{k=0}^{\frac{1}{2}m-2} (-1)^k \left\{\frac{d^{\frac{1}{2}m-k-2}}{dy^{\frac{1}{2}m-k-2}} \mathfrak{G}(y)\right\}_{y=C_\lambda} \left\{\frac{d^k}{dy^k} \omega(y)\right\}_{y=C_\lambda-0}^{y=C_\lambda+0}.$$

(For $C_\lambda = C_1$ or $C_{\frac{1}{2}m}$ the $\frac{d^k}{dy^k} \omega(y)$ at $y = C_1 + 0$ or $C_{\frac{1}{2}m} - 0$, respectively, must obviously be taken to be zero.) Owing to the results of section 5 all terms with $k = 0, 1, \cdots, \frac{m}{2} - 3$ vanish, and the term with $k = \frac{m}{2} - 2$ gives

$$-\sum_{\lambda=1}^{\frac{1}{2}m} (-1)^{\frac{1}{2}m-2} \mathfrak{G}(C_\lambda)(-1)^{\frac{1}{2}m-\lambda} \left(\frac{m}{2} - 1\right)! \frac{1}{\prod_{k=1}^{\lambda-1} (C_k - C_\lambda) \prod_{k=\lambda+1}^{\frac{1}{2}m} (C_\lambda - C_k)}$$

$$= \sum_{\lambda=1}^{\frac{1}{2}m} (-1)^{\lambda-1} \left(\frac{m}{2} - 1\right)! \frac{1}{\prod_{k=1}^{\lambda-1} (C_k - C_\lambda) \prod_{k=\lambda+1}^{\frac{1}{2}m} (C_\lambda - C_k)} \mathfrak{G}(C_\lambda).$$

The second line vanishes, since $\frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}} \omega(y)$ is zero everywhere, as observed above.

Finally

$$\overline{\mathfrak{F}(y)} = \sum_{\lambda=1}^{\frac{1}{2}m} (-1)^{\lambda-1} \left(\frac{m}{2} - 1\right)! \frac{1}{\prod_{k=1}^{\lambda-1} (C_k - C_\lambda) \prod_{k=\lambda+1}^{\frac{1}{2}m} (C_\lambda - C_k)} \mathfrak{G}(C_\lambda).$$

For

$$\mathfrak{B}(z) = \prod_{k=1}^{\frac{1}{2}m} (z - C_k)$$

we have

$$\left\{\frac{d}{dz} \mathfrak{B}(z)\right\}_{z=C_\lambda} = \prod_{k=1}^{\frac{1}{2}m}{}_{(k\neq\lambda)} (C_\lambda - C_k)$$

$$= (-1)^{\lambda-1} \prod_{k=1}^{\lambda-1} (C_k - C_\lambda) \prod_{k=\lambda+1}^{\frac{1}{2}m} (C_\lambda - C_k).$$

Therefore the above formula can also be written

$$\overline{\mathfrak{F}(y)} = \left(\frac{m}{2} - 1\right)! \sum_{\lambda=1}^{\frac{1}{2}m} \frac{\mathfrak{G}(C_\lambda)}{\left\{\frac{d}{dz} \mathfrak{B}(z)\right\}_{z=C_\lambda}}.$$

Observe that the right-hand side of the above formula (which can also be easily expressed in terms of determinants) is a well-known approximate ex-

pression for $\frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}} \mathfrak{G}(y)$, as a (repeated) difference quotient of the values $\mathfrak{G}(C_\lambda)$, $\lambda = 1, \cdots, \frac{m}{2}$. It is therefore very satisfactory that this expression gives the mean of

$$\mathfrak{F}(y) = \frac{d^{\frac{1}{2}m-1}}{dy^{\frac{1}{2}m-1}} \mathfrak{G}(y).$$

Appendix. We return to the normal distribution of $x_1, \cdots, x_n$ as described in section 1, and to the quantities s^2, δ^2, η given there. We denote means with respect to that distribution by $\overline{(\cdots)}$.

It was observed by B. I. Hart and mentioned by J. D. Williams[3] by comparing the known expressions for their moments, that every moment of $\eta = \frac{\delta^2}{s^2}$ is the quotient of the corresponding moments of δ^2 and of s^2. That is

$$\overline{\left(\frac{\delta^{2p}}{s^{2p}}\right)} = \frac{\overline{\delta^{2p}}}{\overline{s^{2p}}}, \qquad (p = 0, 1, 2, \cdots).$$

This indicates some kind of independence relation involving δ^2 and s^2. The considerations which follow are intended to clarify this situation.

The above relation may be written

$$\overline{s^{2p}}\ \overline{\eta^p} = \overline{s^{2p}\eta^p},$$

or, more generally,

$$\overline{s^q}\ \overline{\eta^p} = \overline{s^q\eta^p}.$$

We shall prove this by showing that s and η are statistically independent.

We can, as in section 2, make the mean $\xi = 0$, i.e. obtain the $x_1, \cdots, x_n$ distribution law

$$c^n e^{-\sum_{\mu=1}^{n} x_\mu^2/2\sigma^2} dx_1 \cdots dx_n.$$

And then, again as in section 2, perform a linear orthogonal transformation, carrying $x_1, \cdots, x_n$ into, say $x_1', \cdots, x_n'$ which leaves the distribution law in its original form

$$c^n e^{-\sum_{\mu=1}^{n} x_\mu'^2/2\sigma^2} dx_1' \cdots dx_n',$$

and makes

$$s^2 = \frac{1}{n} \sum_{\mu=1}^{n-1} x_\mu'^2,$$

$$\eta = \frac{n}{n-1} \frac{\sum_{\mu=1}^{n-1} A_\mu x_\mu'^2}{\sum_{\mu=1}^{n-1} x_\mu'^2}.$$

Since x'_n does not occur in s^2, η we must use only the $x'_1, \cdots, x'_{n-1}$ distribution law

$$c^{n-1} e^{-\sum_{\mu=1}^{n-1} x_\mu'^2/2\sigma^2} dx'_1 \cdots dx'_{n-1}.$$

Now we introduce polar coordinates with respect to $x'_1, \cdots, x'_{n-1}$. These consist of a radius r with

$$r^2 = \sum_{\mu=1}^{n-1} x_\mu'^2,$$

and $n - 2$ angular variables $\varphi_1, \cdots, \varphi_{n-2}$, which can be chosen in various ways, and which we need not describe more closely. At any rate

$$dx'_1 \cdots dx'_{n-1} = r^{n-2} dr w(\varphi_1, \cdots, \varphi_{n-2}) \, d\varphi_1 \cdots d\varphi_{n-2}$$

where we need not determine the weight function $w(\varphi_1, \cdots, \varphi_{n-2})$. Consequently the distribution law is

$$c^{n-1} e^{-r^2/2\sigma^2} r^{n-2} dr w(\varphi_1, \cdots, \varphi_{n-2}) \, d\varphi_1 \cdots d\varphi_{n-2}.$$

Thus the coordinate r and the coordinates $\varphi_1, \cdots, \varphi_{n-2}$ are independent of each other.

Next

$$s^2 = \frac{1}{n} r^2,$$

and η is a homogeneous function of $x'_1, \cdots, x'_{n-1}$ of degree zero, i.e. it is independent of r. So s is a function of r alone, and η is a function of $\varphi_1, \cdots, \varphi_{n-2}$ alone. Consequently s and η likewise are independent.

Added in proof:

After this manuscript was completed, Dr. T. Koopmans informed the author of several results of his own, which he obtained in connection with other statistical investigations. They have many points of contact with this investigation, and will appear in the near future in the *Annals of Mathematical Statistics*. The author wishes to express his thanks to Dr. T. Koopmans for his communications.

SOME EXAMPLES OF ASYMPTOTICALLY MOST POWERFUL TESTS

By Abraham Wald[1]

Columbia University

1. Introduction. In a previous paper[2] the author gave the definition of an asymptotically most powerful test and has shown that the commonly used tests, based on the maximum likelihood estimate, are asymptotically most powerful.

In this paper some further examples of asymptotically most powerful tests will be given. Let us first restate the definition of an asymptotically most powerful test. Let $f(x, \theta)$ be the probability density of a variate x involving an unknown parameter θ. For testing the hypothesis $\theta = \theta_0$ by means of n independent observations $x_1, \cdots, x_n$ on x we have to choose a region of rejection W_n in the n-dimensional sample space. Denote by $P(W_n \mid \theta)$ the probability that the sample point $E = (x_1, \cdots, x_n)$ will fall in W_n under the assumption that θ is the true value of the parameter. For any region U_n of the n-dimensional sample space denote by $g(U_n)$ the greatest lower bound of $P(U_n \mid \theta)$. For any pair of regions U_n and T_n denote by $L(U_n, T_n)$ the least upper bound of

$$P(U_n \mid \theta) - P(T_n \mid \theta).$$

In all that follows we shall denote a region of the n-dimensional sample space by a capital letter with the subscript n.

Definition 1: A sequence $\{W_n\}$ ($n = 1, 2, \cdots$, ad inf.) of regions is said to be an asymptotically most powerful test of the hypothesis $\theta = \theta_0$ on the level of significance α if $P(W_n \mid \theta_0) = \alpha$ and if for any sequence $\{Z_n\}$ of regions for which $P(Z_n \mid \theta_0) = \alpha$ the inequality

$$\limsup_{n \to \infty} L(Z_n, W_n) \leq 0$$

holds.

Definition 2: A sequence $\{W_n\}$ ($n = 1, 2, \cdots$, ad inf.) of regions is said to be an asymptotically most powerful unbiased test of the hypothesis $\theta = \theta_0$ on the level of significance α if $P(W_n \mid \theta_0) = \lim_{n=\infty} g(W_n) = \alpha$, and if for any sequence $\{Z_n\}$ of regions for which $P(Z_n \mid \theta_0) = \lim_{n=\infty} g(Z_n) = \alpha$, the inequality

$$\limsup_{n \to \infty} L(Z_n, W_n) \leq 0$$

holds.

---

[1] Research under a grant-in-aid of the Carnegie Corporation of New York.

[2] "Asymptotically most powerful tests of statistical hypotheses," *Annals of Math. Stat.* Vol. 12 (1941).

Consider the expression

$$y_n(\theta) = \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} \frac{\partial}{\partial\theta} \log f(x_\alpha, \theta). \tag{1}$$

Let W'_n be the region defined by the inequality $y_n(\theta_0) \geq c'_n$, W''_n defined by the inequality $y_n(\theta_0) \leq c''_n$, and W_n defined by the inequality $| y_n(\theta_0) | \geq c_n$, where the constants c'_n, c''_n and c_n are chosen such that

$$P(W'_n \mid \theta_0) = P(W''_n \mid \theta_0) = P(W_n \mid \theta_0) = \alpha.$$

It will be shown in this paper that under certain restrictions on the probability density $f(x, \theta)$ the sequence $\{W'_n\}$ is an asymptotically most powerful test of the hypothesis $\theta = \theta_0$ if θ takes only values $\geq \theta_0$. Similarly $\{W''_n\}$ is an asymptotically most powerful test if θ takes only values $\leq \theta_0$. Finally $\{W_n\}$ is an asymptotically most powerful unbiased test if θ can take any real value.

Another example of an asymptotically most powerful unbiased test of the hypothesis $\theta = \theta_0$, as it will be shown, is the critical region of type A in the Neyman-Pearson theory of testing hypotheses. This fact gives a strong justification for the use of the critical region of type A.

2. Assumptions on the density function. Let ω be a subset of the real axis. Denote by θ^* a real variable which takes only values in ω and let θ be a variable which can take any real value. For any function $\psi(x)$ we denote by $E_\theta\psi(x)$ the expected value of $\psi(x)$ under the assumption that θ is the true value of the parameter, i.e.

$$E_\theta \psi(x) = \int_{-\infty}^{+\infty} \psi(x) f(x, \theta)\, dx.$$

For any x, for any positive δ and for any real value θ_1 denote by $\varphi_1(x, \theta_1, \delta)$ the greatest lower bound, and by $\varphi_2(x, \theta_1, \delta)$ the least upper bound of $\dfrac{\partial^2}{\partial\theta^2} \log f(x, \theta)$ in the interval $\theta_1 - \delta \leq \theta \leq \theta_1 + \delta$. In all that follows the symbol θ_i^*, for any integer i, will denote a value of θ^*, i.e., θ_i^* is a point of ω.

We say that a value θ lies in the ϵ-neighborhood of ω if there exists a value θ^* such that $| \theta - \theta^* | \leq \epsilon$.

Throughout the paper the following assumptions on $f(x, \theta)$ will be made:

Assumption 1: *For any pair of sequences $\{\theta_n\}$ and $\{\theta_n^*\}$ $(n = 1, 2, \cdots, \text{ad inf.})$ for which*

$$\lim_{n=\infty} E_{\theta_n} \frac{\partial}{\partial\theta} \log f(x, \theta_n^*) = 0$$

also

$$\lim_{n=\infty} (\theta_n - \theta_n^*) = 0.$$

Furthermore there exists a positive ϵ *such that* $E_\theta\left[\frac{\partial}{\partial\theta}\log f(x, \theta_1)\right]^2$ *is a bounded function of* θ *and* θ_1, $E_\theta \frac{\partial}{\partial\theta}\log f(x, \theta_1)$ *is a continuous function of* θ *and* θ_1 *and* $E_{\theta_1}\left[\frac{\partial}{\partial\theta}\log f(x, \theta_1)\right]^2 = d(\theta_1)$ *has a positive lower bound, where* θ_1 *can take any value in the* ϵ*-neighborhood of* ω.

ASSUMPTION 2: *There exists a positive* k_0 *such that* $E_{\theta_2}\varphi_1(x, \theta_1, \delta)$ *and* $E_{\theta_2}\varphi_2(x, \theta_1, \delta)$ *are uniformly continuous functions in the domain D defined as follows: the variables* θ_1 *and* θ_2 *may take any value in the* k_0*-neighborhood of* ω *and* δ *may take any value for which* $|\delta| \leq k_0$. *Furthermore it is assumed that*

$$E_{\theta_2}[\varphi_i(x, \theta_1, \delta)]^2, \qquad (i = 1, 2)$$

are bounded functions of θ_1, θ_2 *and* δ *in D.*

ASSUMPTION 3: *There exists a positive* k_0 *such that*

$$\int_{-\infty}^{+\infty} \frac{\partial}{\partial\theta} f(x, \theta)\, dx = \int_{-\infty}^{+\infty} \frac{\partial^2}{\partial\theta^2} f(x, \theta)\, dx = 0$$

for all θ *in the* k_0*-neighborhood of* ω.

Assumption 3 means simply that we may differentiate with respect to θ under the integral sign. In fact,

$$\int_{-\infty}^{+\infty} f(x, \theta)\, dx = 1,$$

identically in θ. Hence

$$\frac{\partial}{\partial\theta}\int_{-\infty}^{+\infty} f(x, \theta)\, dx = \frac{\partial^2}{\partial\theta^2}\int_{-\infty}^{+\infty} f(x, \theta)\, dx = 0.$$

Differentiating under the integral sign we obtain the relations in Assumption 3.

ASSUMPTION 4: *There exists a positive* k_0 *and a positive* η *such that*

$$E_\theta\left[\frac{\partial}{\partial\theta}\log f(x, \theta)\right]^{2+\eta}$$

is a bounded function of θ *in the* k_0*-neighborhood of* ω.

3. Some propositions. PROPOSITION 1: *To any positive* β *there exists a positive* γ *such that*

$$\lim_{n=\infty} P\left\{\frac{1}{\sqrt{n}}\,|y_n(\theta^*)| > \gamma \mid \theta\right\} = 1$$

uniformly in θ^* *and for all* θ *for which* $|\theta - \theta^*| \geq \beta$.

PROOF: From Assumption 1 it follows that $\left| E_\theta \frac{\partial}{\partial\theta}\log f(x, \theta^*)\right|$ has a positive

lower bound in the domain $|\theta - \theta^*| \geq \beta$. Since according to Assumption 1 $E_\theta\left[\frac{\partial}{\partial\theta}\log f(x, \theta^*)\right]^2$ is a bounded function of θ and θ^*, Proposition 1 easily follows.

PROPOSITION 2: *There exists a positive ϵ such that*

$$\lim_{n=\infty} P[y_n(\theta) < t \mid \theta] = N(t \mid \theta)$$

uniformly in t and for all θ in the ϵ-neighborhood of ω where

$$d(\theta) = -E_\theta \frac{\partial^2}{\partial\theta^2}\log f(x, \theta) = E_\theta\left[\frac{\partial}{\partial\theta}\log f(x, \theta)\right]^2 \tag{2}$$

and

$$N(t \mid \theta) = \frac{1}{\sqrt{2\pi d(\theta)}}\int_{-\infty}^{t} e^{-\frac{1}{2}v^2/d(\theta)}\,dv. \tag{3}$$

Proposition 2 follows easily from Assumptions 3 and 4 and the general limit theorems.

PROPOSITION 3: *There exists a positive ϵ such that for any bounded sequence $\{\mu_n\}$*

$$\lim_{n=\infty}\left\{P\left[y_n(\theta) < t \mid \theta + \frac{\mu_n}{\sqrt{n}}\right] - \int_{-\infty}^{t} e^{\mu_n v - \frac{1}{2}\mu_n^2 d(\theta)}\,dN(v \mid \theta)\right\} = 0$$

uniformly in t and for all θ in the ϵ-neighborhood of ω.

PROOF: We have

$$y_n\left(\theta + \frac{\mu_n}{\sqrt{n}}\right) = y_n(\theta) + \frac{\mu_n}{\sqrt{n}}\frac{1}{\sqrt{n}}\sum_\alpha \frac{\partial^2}{\partial\theta^2}\log f(x_\alpha, \theta_n') \tag{4}$$

where θ_n' lies in the interval $\left[\theta, \theta + \frac{\mu_n}{\sqrt{n}}\right]$. From Assumption 2 and the above equation we easily obtain

$$\lim_{n=\infty}\left\{P\left[y_n\left(\theta + \frac{\mu_n}{\sqrt{n}}\right) < t \mid \theta + \frac{\mu_n}{\sqrt{n}}\right] - P\left[y_n(\theta) - \mu_n d(\theta) < t \mid \theta + \frac{\mu_n}{\sqrt{n}}\right]\right\} = 0 \tag{5}$$

uniformly in t and for all θ in the ϵ-neighborhood of ω. From Proposition 2 and (5) we get

$$\lim_{n=\infty}\left\{\int_{-\infty}^{t} dN(v \mid \theta) - P\left[y_n(\theta) < t + \mu_n d(\theta) \,\middle|\, \theta + \frac{\mu_n}{\sqrt{n}}\right]\right\} = 0$$

or

$$\lim_{n=\infty}\left\{\int_{-\infty}^{t-\mu_n d(\theta)} dN(v \mid \theta) - P\left[y_n(\theta) < t \mid \theta + \frac{\mu_n}{\sqrt{n}}\right]\right\} = 0 \tag{6}$$

uniformly in t and for all θ in the ϵ-neighbourhood of ω. This proves Proposition 3.

PROPOSITION 4: *There exists a positive ϵ such that for any positive γ and for any sequence $\{\mu_n\}$ for which $\lim_{n=\infty} |\mu_n| = \infty$*

$$\lim_{n=\infty} P\left\{|y_n(\theta^*)| > \gamma \mid \theta^* + \frac{\mu_n}{\sqrt{n}}\right\} = 1$$

uniformly in θ^*.

PROOF: If there exists a positive β such that $\left|\frac{\mu_n}{\sqrt{n}}\right| > \beta$ for almost all n, Proposition 4 follows from Proposition 1. Hence we have to consider only the case $\lim_{n=\infty} \frac{\mu_n}{\sqrt{n}} = 0$. Since

$$E_{\theta^*+(\mu_n/\sqrt{n})}\, y_n\left(\theta^* + \frac{\mu_n}{\sqrt{n}}\right) = 0,$$

we get from (4)

$$E_{\theta^*+(\mu_n/\sqrt{n})}[y_n(\theta^*)] + \mu_n E_{\theta^*+(\mu_n/\sqrt{n})} \frac{\sum_\alpha \frac{\partial^2}{\partial\theta^2} \log f(x_\alpha, \theta_n')}{n} = 0. \tag{7}$$

Since $\lim \frac{\mu_n}{\sqrt{n}} = 0$, we have on account of Assumption 2

$$\lim_{n=\infty} E_{\theta^*+(\mu_n/\sqrt{n})} \frac{\sum_\alpha \frac{\partial^2}{\partial\theta^2} \log f(x_\alpha, \theta_n')}{n} = E_{\theta^*} \frac{\partial^2}{\partial\theta^2} \log f(x, \theta^*)$$
$$= -E_{\theta^*}\left[\frac{\partial}{\partial\theta} \log f(x, \theta^*)\right]^2 = -d(\theta^*)$$

uniformly in θ^*. According to Assumption 1 $d(\theta^*)$ has a positive lower bound; hence on account of $\lim |\mu_n| = \infty$ we obtain from (7)

$$\lim_{n=\infty} |E_{\theta^*+(\mu_n/\sqrt{n})}\, y_n(\theta^*)| = \infty \tag{8}$$

uniformly in θ^*. The variance of $y_n(\theta^*)$ is equal to the variance of $\frac{\partial}{\partial\theta} \log f(x, \theta^*)$. On account of Assumption 1 the variance of $\frac{\partial}{\partial\theta} \log f(x, \theta^*)$ (under the assumption that $\theta^* + \frac{\mu_n}{\sqrt{n}}$ is the true value of the parameter) is a bounded function. Hence Proposition 4 is proved on account of (8).

PROPOSITION 5: *Let $\{W_n(\theta^*)\}$ be a sequence of regions of size α, i.e. $P[W_n(\theta^*) \mid \theta^*] = \alpha$, and let $V_n(\theta^*, y)$ be the region defined by the inequality*

$y_n(\theta^*) < y$. Let $U_n(\theta^*, y)$ be the intersection of $V_n(\theta^*, y)$ and $W_n(\theta^*)$ and denote $P[U_n(\theta^*, y) \mid \theta^*]$ by $F_n(y \mid \theta^*)$. Denote furthermore $P\left[W_n(\theta^*) \mid \theta^* + \frac{\mu}{\sqrt{n}}\right]$ by $G(\theta^*, \mu, n)$. If $\{\theta_n^*\}$ and $\{\mu_n\}$ are two sequences such that $\lim_{n=\infty} d(\theta_n^*) = d$; $\lim_{n=\infty} F_n(y \mid \theta_n^*) = F(y)$ and $\lim \mu_n = \mu$ then

$$\lim_{n=\infty} G(\theta_n^*, \mu_n, n) = \int_{-\infty}^{+\infty} e^{\mu y - \frac{1}{2}\mu^2 d}\, dF(y).$$

PROOF: Let $\lim \mu_n = \mu$ and consider the Taylor expansion

$$\begin{aligned}\sum_\alpha \log f\left(x_\alpha, \theta^* + \frac{\mu_n}{\sqrt{n}}\right) &= \sum_\alpha \log f(x_\alpha, \theta^*) + \frac{\mu_n}{\sqrt{n}} \sum_\alpha \frac{\partial}{\partial\theta} \log f(x_\alpha, \theta^*) \\ &\quad + \frac{1}{2}\frac{\mu_n^2}{n}\frac{\partial^2}{\partial\theta^2}\sum_\alpha \log f(x_\alpha, \theta_n')\end{aligned} \tag{9}$$

where θ_n' lies in the interval $\left[\theta^*, \theta^* + \frac{\mu_n}{\sqrt{n}}\right]$. From this we easily get on account of Assumption 2 and the fact that $\{\mu_n\}$ is bounded

$$\log \prod_{\alpha=1}^{n} \frac{f\left(x_\alpha, \theta^* + \frac{\mu_n}{\sqrt{n}}\right)}{f(x_\alpha, \theta^*)} = \mu_n y_n(\theta^*) - \tfrac{1}{2}\mu_n^2\, d(\theta^*) + \epsilon(\theta^*, n) \tag{10}$$

where for arbitrary positive η

$$\lim_{n=\infty} P\left\{\mid \epsilon(\theta^*, n) \mid < \eta \mid \theta^* + \frac{\mu_n}{\sqrt{n}}\right\} = 1 \tag{11}$$

uniformly in θ^*. Denote by $R_n(\theta^*)$ the region defined by

$$\mid \epsilon(\theta^*, n) \mid < \eta > 0. \tag{12}$$

On account of (11) we have

$$\lim_{n=\infty} P\left[R_n(\theta^*) \mid \theta^* + \frac{\mu_n}{\sqrt{n}}\right] = 1, \tag{13}$$

uniformly in θ^*. Denote the intersection of $R_n(\theta^*)$ and $W_n(\theta^*)$ by $Q_n(\theta^*)$, and the intersection of $R_n(\theta^*)$ and $U_n(\theta^*, y)$ by $T_n(\theta^*, y)$. Furthermore denote $P[T_n(\theta^*, y) \mid \theta^*]$ by $\bar{F}_n(y \mid \theta^*)$. Then we have

$$\begin{aligned} e^{-\eta}\int_{-\infty}^{t} e^{\mu_n y - \frac{1}{2}\mu_n^2 d(\theta^*)}\, d\bar{F}_n(y \mid \theta^*) &\le P\left[T_n(\theta^*, t) \mid \theta^* + \frac{\mu_n}{\sqrt{n}}\right] \\ &\le e^{\eta}\int_{-\infty}^{t} e^{\mu_n y - \frac{1}{2}\mu_n^2 d(\theta^*)}\, d\bar{F}_n(y \mid \theta^*)\end{aligned} \tag{14}$$

for all values of t and θ^*. Furthermore we obviously have

$$\lim_{n=\infty} \left\{ G(\theta^*, \mu_n, n) - P\left[Q_n(\theta^*) \mid \theta^* + \frac{\mu_n}{\sqrt{n}} \right] \right\} = 0 \tag{15}$$

uniformly in θ^*, and

$$\lim_{n=\infty} [\bar{F}_n(t \mid \theta^*) - F_n(t \mid \theta^*)] = 0 \tag{16}$$

uniformly in θ^* and t. Since η may be chosen arbitrarily small, it follows from (14) and (15) that to any $\epsilon > 0$, η may be chosen such that

$$\limsup_{n=\infty} \left| G(\theta_n^*, \mu_n, n) - \int_{-\infty}^{+\infty} e^{\mu_n t - \frac{1}{2}\mu_n^2 d(\theta_n^*)} \, d\bar{F}_n(t \mid \theta_n^*) \right| \leqq \frac{\epsilon}{2} \tag{17}$$

for any sequence $\{\theta_n^*\}$.

To each ϵ let L_ϵ be a positive number such that L_ϵ depends only on ϵ and

$$\int_{-\infty}^{-L_\epsilon} e^{\mu_n t - \frac{1}{2}\mu_n^2 d(\theta^*)} \, dN(t \mid \theta^*) + \int_{L_\epsilon}^{\infty} e^{\mu_n t - \frac{1}{2}\mu_n^2 d(\theta^*)} \, dN(t \mid \theta^*) \leq \frac{\epsilon}{2} \tag{18}$$

for all n and for all values of θ^*. Since $d(\theta^*)$ has a positive lower and a finite upper bound, it is easy to verify that such a L_ϵ exists. From (18) and Proposition 3 it follows

$$\begin{aligned} \limsup_{n \to \infty} \Bigg\{ & P\left[y_n(\theta_n^*) < -L_\epsilon \mid \theta_n^* + \frac{\mu_n}{\sqrt{n}} \right] \\ & + P\left[y_n(\theta_n^*) > L_\epsilon \mid \theta_n^* + \frac{\mu_n}{\sqrt{n}} \right] \Bigg\} \leq \frac{\epsilon}{2} \end{aligned} \tag{19}$$

for any arbitrary sequence $\{\theta_n^*\}$. Since the difference $U_n(\theta^*, t_2) - U_n(\theta^*, t_1)$ is a subset of the difference $V_n(\theta^*, t_2) - V_n(\theta^*, t_1)$ and since $T_n(\theta^*, t_2) - T_n(\theta^*, t_1)$ is a subset of $U_n(\theta^*, t_2) - U_n(\theta^*, t_1)$ for $t_2 > t_1$, we get from (18) and (19)

$$\begin{aligned} \limsup_{n \to \infty} \Bigg\{ & P\left[U_n(\theta_n^*, -L_\epsilon) \mid \theta_n^* + \frac{\mu_n}{\sqrt{n}} \right] + P\left[W_n(\theta_n^*) \mid \theta_n^* + \frac{\mu_n}{\sqrt{n}} \right] \\ & - P\left[U_n(\theta_n^*, L_\epsilon) \mid \theta_n^* + \frac{\mu_n}{\sqrt{n}} \right] \Bigg\} \leq \frac{\epsilon}{2} \end{aligned} \tag{20}$$

and

$$\begin{aligned} \limsup_{n \to \infty} \Bigg\{ & P\left[T_n(\theta_n^*, -L_\epsilon) \mid \theta_n^* + \frac{\mu_n}{\sqrt{n}} \right] + P\left[Q_n(\theta_n^*) \mid \theta_n^* + \frac{\mu_n}{\sqrt{n}} \right] \\ & - P\left[T_n(\theta_n^*, L_\epsilon) \mid \theta_n^* + \frac{\mu_n}{\sqrt{n}} \right] \Bigg\} \leq \frac{\epsilon}{2} \end{aligned} \tag{21}$$

for any sequence $\{\theta_n^*\}$. On account of (14) we get from (21)

$$e^{-\eta} \limsup_{n \to \infty} \left\{ \int_{-\infty}^{-L_\epsilon} e^{\mu_n t - \frac{1}{2}\mu_n^2 d(\theta_n^*)} \, d\bar{F}_n(t \mid \theta_n^*) + \int_{L_\epsilon}^{\infty} e^{\mu_n t - \frac{1}{2}\mu_n^2 d(\theta_n^*)} \, d\bar{F}_n(t \mid \theta_n^*) \right\} \leq \frac{\epsilon}{2}. \tag{22}$$

From (17) and (22) we obtain

$$(23)\qquad \limsup_{n\to\infty}\left| G(\theta_n^*, \mu_n, n) - \int_{-L_\epsilon}^{L_\epsilon} e^{\mu_n t - \frac{1}{2}\mu_n^2 d(\theta_n^*)}\, d\bar{F}_n(t \mid \theta_n^*)\right| \leq \epsilon\left(\frac{1+e^\eta}{2}\right)$$

for any sequence $\{\theta_n^*\}$. Consider now the sequence $\{\theta_n^*\}$ which satisfies the conditions of Proposition 5. Since $F_n(t \mid \theta_n^*)$ converges to $F(t)$ uniformly in t, on account of (16) also $\bar{F}_n(t \mid \theta_n^*)$ converges to $F(t)$ uniformly in t. Hence we obtain from (23)

$$(24)\qquad \limsup_{n\to\infty}\left| G(\theta_n^*, \mu_n, n) - \int_{-L_\epsilon}^{L_\epsilon} e^{\mu t - \frac{1}{2}\mu^2 d}\, dF(t)\right| \leq \epsilon\left(\frac{1+e^\eta}{2}\right).$$

Since ϵ and η may be chosen arbitrarily small, Proposition 5 follows from (24).

4. Some theorems and corollaries. THEOREM 1. *Denote by $S_n(\theta^*)$ the region defined by the inequality $y_n(\theta^*) \geq A_n(\theta^*)$ where $A_n(\theta^*)$ is chosen such that $P[S_n(\theta^*) \mid \theta^*] = \alpha$. For any region $W_n(\theta^*)$ denote by $L_n[W_n(\theta^*)]$ the least upper bound of $P[W_n(\theta^*) \mid \theta] - P[S_n(\theta^*) \mid \theta]$ with respect to θ^* and θ, where θ is restricted to values $\geq \theta^*$. Then for any sequence $\{W_n(\theta^*)\}$ for which $P[W_n(\theta^*) \mid \theta^*] = \alpha$,*

$$\limsup_{n\to\infty} L_n[W_n(\theta^*)] \leq 0.$$

PROOF: Assume that Theorem 1 is not true. Then there exists a sequence of integers $\{n'\}$, a sequence $\{\theta_{n'}^*\}$ and a sequence $\{\theta_{n'}\}$ $(\theta_{n'} \geq \theta_{n'}^*)$ such that

$$(25)\qquad \lim_{n=\infty}\{P[W_{n'}(\theta_{n'}^*) \mid \theta_{n'}] - P[S_{n'}(\theta_{n'}^*) \mid \theta_{n'}]\} = \delta > 0.$$

On account of Proposition 2 and Assumption 2 the sequence $\{A_{n'}(\theta_{n'}^*)\}$ is bounded. Then it follows easily from (25) and Proposition 4 (taking in account that $E_\theta \dfrac{\partial}{\partial\theta}\log f(x, \theta^*) > 0$ for $\theta > \theta^*$

$$(26)\qquad (\theta_{n'} - \theta_{n'}^*)\sqrt{n'} = \mu_{n'} > 0$$

must be bounded. Denote by $\{n''\}$ a subsequence of $\{n'\}$ such that

$$(27)\qquad \lim d(\theta_{n''}^*) = d$$

$$(28)\qquad \lim \mu_{n''} = \mu, \quad \text{and}$$

$$(29)\qquad \lim F_{n''}(t \mid \theta_{n''}^*) = F(t)$$

uniformly in t where

$$F_n(t \mid \theta^*) = P[U_n(\theta^*, t) \mid \theta^*]$$

and $U_n(\theta^*, t)$ is the intersection of $W_n(\theta^*)$ and the region $y_n(\theta^*) < t$. The existence of a subsequence $\{n''\}$ such that (29) holds follows from the fact that

$$(30)\qquad F_n(t_2 \mid \theta^*) - F_n(t_1 \mid \theta^*) \leq \Phi_n(t_2 \mid \theta^*) - \Phi_n(t_1 \mid \theta^*) \quad \text{for} \quad t_2 > t_1,$$

and

$$\lim_{n=\infty} \Phi_n(t \mid \theta_n^*,,) = \frac{1}{\sqrt{2\pi d}} \int_{-\infty}^{t} e^{-\frac{1}{2}v^2/d}\, dv = N(t), \tag{31}$$

where $\Phi_n(t \mid \theta^*)$ denotes the probability $P[y_n(\theta^*) < t \mid \theta^*]$. Furthermore it can easily be shown that

$$\int_{-\infty}^{+\infty} dF(t) = \alpha. \tag{32}$$

On account of Proposition 5 we get from (25), (27), (28), (29), (30) and (31)

$$\int_{-\infty}^{+\infty} e^{\mu t - \frac{1}{2}\mu^2 d}\, dF(t) - \int_{A}^{\infty} e^{\mu t - \frac{1}{2}\mu^2 d}\, dN(t) = \delta, \tag{33}$$

where A denotes a value such that

$$\int_{A}^{\infty} dN(t) = \alpha.$$

It has been shown in a previous paper[3] that (33) leads to a contradiction. Hence Theorem 1 is proved.

THEOREM 2: *Denote by $S_n(\theta^*)$ the region defined by the inequality $y_n(\theta^*) \leq A_n(\theta^*)$ where $A_n(\theta^*)$ is chosen such that $P[S_n(\theta^*) \mid \theta^*] = \alpha$. For any region $W_n(\theta^*)$ denote by $L_n[W_n(\theta^*)]$ the least upper bound of*

$$P[W_n(\theta^*) \mid \theta] - P[S_n(\theta^*) \mid \theta]$$

with respect to θ^* and θ, where θ is restricted to values $\leq \theta^*$. Then for any sequence $\{W_n(\theta^*)\}$ for which $P[W_n(\theta^*) \mid \theta^*) = \alpha$,

$$\limsup_{n \to \infty} L_n[W_n(\theta^*)] \leq 0.$$

The proof is omitted, since it is analogous to that of Theorem 1.

THEOREM 3: *Let $\{W_n(\theta^*)\}$ be for each θ^* a sequence of regions for which $P[W_n(\theta^*) \mid \theta^*] = \alpha$ and $\lim_{n=\infty} g[W_n(\theta^*)] = \alpha$ uniformly in θ^*. Denote by $L_n[W_n(\theta^*)]$ the least upper bound of*

$$P[W_n(\theta^*) \mid \theta] - P[| y_n(\theta^*) | \geq A_n(\theta^*) \mid \theta]$$

with respect to θ and θ^*, where $A_n(\theta^*)$ is chosen such that

$$P[| y_n(\theta^*) | \geq A_n(\theta^*) \mid \theta^*] = \alpha.$$

Then

$$\limsup_{n \to \infty} L_n[W_n(\theta^*)] \leq 0.$$

---

[3] See p. 12 of the paper cited in [2].

PROOF: Denote $P[y_n(\theta^*) < t \mid \theta^*]$ by $\Phi_n(t \mid \theta^*)$ and denote by $F_n(t \mid \theta^*)$ the probability (under the hypothesis $\theta = \theta^*$) of the intersection of $W_n(\theta^*)$ with the region $y_n(\theta^*) < t$. Assume that Theorem 3 is not true. Then there exists a subsequence $\{n''\}$, a sequence $\{\theta^*_{n''}\}$ and a sequence $\{\theta_{n''}\}$ such that

$$\lim_{n=\infty} d(\theta^*_{n''}) = d; \qquad \lim\, (\theta_{n''} - \theta^*_{n''})\sqrt{n''} = \lim \mu_{n''} = \mu;$$

$$\lim F_{n''}(t \mid \theta^*_{n''}) = F(t)$$

uniformly in t, and

$$\int_{-\infty}^{+\infty} e^{\mu t - \frac{1}{2}\mu^2 d}\, dF(t) - \int_{-\infty}^{-A} e^{\mu t - \frac{1}{2}\mu^2 d}\, dN(t) - \int_{A}^{\infty} e^{\mu t - \frac{1}{2}\mu^2 d}\, dN(t) = \delta \tag{34}$$

where A is a positive number such that

$$\int_{-\infty}^{-A} dN(t) = \frac{\alpha}{2}, \qquad \text{and} \qquad N(t) = \frac{1}{\sqrt{2\pi d}} \int_{-\infty}^{t} e^{-\frac{1}{2}v^2/d}\, dv.$$

This can be proved in the same way as (33) has been proved. The author has shown in a previous paper[4] that (34) leads to a contradiction. Hence Theorem 3 is proved.

THEOREM 4: *Denote by $A_n(\theta^*)$ the region of type[5] A of size α for testing the hypothesis $\theta = \theta^*$. Denote by $B_n(\theta^*)$ the region $| y_n(\theta^*) | \geq C_n(\theta^*)$ where $C_n(\theta^*)$ is determined such that*

$$P[| y_n(\theta^*) | \geq C_n(\theta^*) \mid \theta^*] = \alpha.$$

Then, under the assumption that $E_\theta \left[\frac{\partial^2}{\partial\theta^2} \log f(x, \theta^*) \right]^2$ *is bounded,*

$$\lim_{n=\infty} \{P[A_n(\theta^*) \mid \theta] - P[B_n(\theta^*) \mid \theta]\} = 0$$

uniformly in θ and θ^*.

PROOF: The region $A_n(\theta^*)$ is given by the inequality[6]

$$\begin{aligned} &\left[\sum_\alpha \frac{\partial}{\partial\theta} \log f(x_\alpha, \theta^*)\right]^2 \\ &\quad + \sum_\alpha \frac{\partial^2}{\partial\theta^2} \log f(x_\alpha, \theta^*) \geq k'_n(\theta^*) \left[\sum_\alpha \frac{\partial}{\partial\theta} \log f(x_\alpha, \theta^*)\right] + k''_n(\theta^*), \end{aligned} \tag{35}$$

where $k'_n(\theta^*)$ and $k''_n(\theta^*)$ are chosen such that $A_n(\theta^*)$ should be unbiased and of size α. The inequality (35) can be written also in the form

$$[y_n(\theta^*)]^2 + \frac{1}{n} \sum_\alpha \frac{\partial^2}{\partial\theta^2} \log f(x_\alpha, \theta^*) \geq l'_n(\theta^*) y_n(\theta^*) + l''_n(\theta^*). \tag{36}$$

---

[4] See p. 14 of the paper cited in [2].

[5] Neyman, J. and Pearson, E. S., "Contributions to the theory of testing statistical hypotheses," *Stat. Res. Mem.*, Vol. 1.

[6] See the paper cited in [5].

Let $\{\mu_n\}$ be a bounded sequence. From Assumption 2 it follows that for any positive ϵ

$$(37)\qquad P\left\{\left|\frac{1}{n}\sum_{\alpha}\frac{\partial^2}{\partial\theta^2}\log f(x_\alpha,\theta^*)+d(\theta^*)\right|<\epsilon\mid\theta^*+\frac{\mu_n}{\sqrt{n}}\right\}=1$$

uniformly in θ^*. Since (37) holds for arbitrarily small ϵ, we get easily on account of Proposition 3

$$(38)\qquad \lim_{n=\infty}\left\{P\left[A_n(\theta^*)\mid\theta^*+\frac{\mu_n}{\sqrt{n}}\right]-P\left[A'_n(\theta^*)\mid\theta^*+\frac{\mu_n}{\sqrt{n}}\right]\right\}=0$$

uniformly in θ^*, where $A'_n(\theta^*)$ is defined by

$$(39)\qquad [y_n(\theta^*)]^2\geq l'_n(\theta^*)y_n(\theta^*)+l''_n(\theta^*)+d(\theta^*).$$

Since $A_n(\theta^*)$ is unbiased and of size α, we have on account of (38) and (39)

$$(40)\qquad \lim l'_n(\theta^*)=0 \quad\text{and}$$

$$(41)\qquad \lim l''_n(\theta^*)+d(\theta^*)=\lambda(\theta^*)>0$$

uniformly in θ^*, where $\lambda(\theta^*)$ is given by the condition

$$(42)\qquad \frac{1}{\sqrt{2\pi d(\theta^*)}}\int_{-\sqrt{\lambda(\theta^*)}}^{+\sqrt{\lambda(\theta^*)}}e^{-\frac{1}{2}t^2/d(\theta^*)}\,dt=\alpha.$$

Inequality (39) is obviously equivalent to the simultaneous inequalities:

$$y_n(\theta^*)\leq c'_n(\theta^*)\qquad\text{and}\qquad y_n(\theta^*)\geq c''_n(\theta^*)$$

where $c'_n(\theta^*)$ and $c''_n(\theta^*)$ are the roots of the equation in $y_n(\theta^*)$

$$[y_n(\theta^*)]^2=l'_n(\theta^*)y_n(\theta^*)+l''_n(\theta^*)+d(\theta^*).$$

Since

$$\lim c'_n(\theta^*)=-\sqrt{\lambda(\theta^*)}\qquad\text{and}\qquad\lim c''_n(\theta^*)=+\sqrt{\lambda(\theta^*)}$$

uniformly in θ^*, from Proposition 3 it follows that

$$(43)\qquad \lim_{n=\infty}\left\{P\left[A_n(\theta^*)\mid\theta^*+\frac{\mu_n}{\sqrt{n}}\right]-\int_{-\infty}^{-\sqrt{\lambda(\theta^*)}}e^{\mu_n t-\frac{1}{2}\mu_n^2 d(\theta^*)}\,dN(t\mid\theta^*)-\int_{+\sqrt{\lambda(\theta^*)}}^{\infty}e^{\mu_n t-\frac{1}{2}\mu_n^2 d(\theta^*)}\,dN(t\mid\theta^*)\right\}=0$$

uniformly in θ^*.

Now let us consider a sequence $\{\nu_n\}$ such that $\lim|\nu_n|=\infty$ and $\lim\dfrac{\nu_n}{\sqrt{n}}=0$.

We shall prove that

$$(44) \qquad P\left[A_n(\theta^*) \mid \theta^* + \frac{\nu_n}{\sqrt{n}}\right] = 1$$

uniformly in θ^*. Since $E_\theta\left[\frac{\partial^2}{\partial\theta^2}\log f(x, \theta^*)\right]^2$ is assumed to be bounded,

$$(45) \qquad E_{\theta^*+(\nu_n/\sqrt{n})}\left[\frac{\partial^2}{\partial\theta^2}\log f(x, \theta^*)\right]$$

and

$$(46) \qquad E_{\theta^*+(\nu_n/\sqrt{n})}\left[\frac{\partial^2}{\partial\theta^2}\log f(x, \theta^*)\right]^2$$

are bounded functions of θ^* and n. We get by Taylor expansion

$$(47) \qquad \begin{aligned}\sum_\alpha \frac{\partial}{\partial\theta}\log f(x_\alpha, \theta^*) = \sum_\alpha \frac{\partial}{\partial\theta}\log f\left(x_\alpha, \theta^* + \frac{\nu_n}{\sqrt{n}}\right) \\ - \frac{\nu_n}{\sqrt{n}}\sum_\alpha \frac{\partial^2}{\partial\theta^2}\log f(x_\alpha, \bar{\theta}_n^*)\end{aligned}$$

where $\bar{\theta}_n^*$ lies in $\left[\theta^*, \theta^* + \frac{\nu_n}{\sqrt{n}}\right]$. Hence

$$(48) \qquad E_{\theta^*+(\nu_n/\sqrt{n})}[y_n(\theta^*)] = -\nu_n E_{\theta^*+(\nu_n/\sqrt{n})}\left[\frac{1}{n}\frac{\partial^2}{\partial\theta^2}\sum_\alpha \log f(x_\alpha, \bar{\theta}_n^*)\right].$$

From Assumption 2 and $\lim |\nu_n| = \infty$ it follows that the absolute value of the right hand side of (48) converges to ∞. Hence

$$\lim | E_{\theta^*+\nu_n/\sqrt{n}}[y_n(\theta^*)] | = \infty.$$

Since on account of Assumption 1

$$E_{\theta^*+(\nu_n/\sqrt{n})}\left[\frac{\partial}{\partial\theta}\log f(x_\alpha, \theta^*)\right]^2$$

is a bounded function of n and θ^*, also the variance of $y_n(\theta^*)$ (under the assumption that $\theta = \theta^* + \nu_n/\sqrt{n}$ is the true value of the parameter) is a bounded function of n and θ^*. Hence for any arbitrary large constant C

$$(49) \qquad \lim P\left[\,| y_n(\theta^*) | \geq C \mid \theta^* + \frac{\nu_n}{\sqrt{n}}\right] = 1,$$

uniformly in θ^*. The equation (44) follows easily from (36), (40), (41), (45), (46) and (49).

Consider a sequence $\{\rho_n\}$ such that $\left|\frac{\rho_n}{\sqrt{n}}\right| > \beta > 0$ for all n. Then it follows easily from Proposition 1 that for any arbitrary C

$$\lim P\left[\,|\, y_n(\theta^*) \,|\, \geq C \mid \theta^* + \frac{\rho_n}{\sqrt{n}}\right] = 1 \tag{50}$$

uniformly in θ^*. Since $E_\theta\left[\frac{\partial^2}{\partial\theta^2}\log f(x_\alpha\,,\,\theta^*)\right]^2$ is assumed to be bounded, and therefore also $E_\theta\frac{\partial^2}{\partial\theta^2}\log f(x,\theta^*)$ is bounded, there exists a finite g such that

$$\lim P\left\{\left|\frac{1}{n}\sum_\alpha \frac{\partial^2}{\partial\theta^2}\log f(x_\alpha\,,\,\theta^*)\right| < g \mid \theta^* + \frac{\rho_n}{\sqrt{n}}\right\} = 1 \tag{51}$$

uniformly in θ^*. From (36), (40), (41), (50) and (51) it follows

$$\lim P\left[A_n(\theta^*) \mid \theta^* + \frac{\rho_n}{\sqrt{n}}\right] = 1 \tag{52}$$

uniformly in θ^*. Since on account of Propositions 3 and 4, the relations (43), (44) and (52) hold if we substitute $B_n(\theta^*)$ for $A_n(\theta^*)$, Theorem 4 is proved.

If Assumptions 1–4 are fulfilled for the set ω consisting of the single point $\theta = \theta_0$, then we get from Theorems 1–4 the following corollaries:

COROLLARY 1: *Let W_n' be the region defined by the inequality $y_n(\theta_0) \geq c_n'$, W_n'' defined by the inequality $y_n(\theta_0) \leq c_n''$, and W_n defined by the inequality $|\, y_n(\theta_0)\,| \geq c_n$, where the constants c_n', c_n'' and c_n are chosen such that*

$$P(W_n' \mid \theta_0) = P(W_n'' \mid \theta_0) = P(W_n \mid \theta_0) = \alpha.$$

Then $\{W_n'\}$ is an asymptotically most powerful test of the hypothesis $\theta = \theta_0$ if θ takes only values $\geq \theta_0$. Similarly $\{W_n''\}$ is an asymptotically most powerful test if θ takes only values $\leq \theta_0$. Finally $\{W_n\}$ is an asymptotically most powerful unbiased test if θ can take any real value.

COROLLARY 2: *The sequence $\{A_n(\theta_0)\}$ is an asymptotically most powerful unbiased test of the hypothesis $\theta = \theta_0$, where $A_n(\theta_0)$ denotes the critical region of type A for testing $\theta = \theta_0$.*

ON THE DISTRIBUTION OF THE QUOTIENT OF TWO CHANCE VARIABLES

By J. H. Curtiss

Cornell University

1. **Introduction.** Although the quotient of two chance variables appears frequently in mathematical statistics, the methods used in the literature to derive the distributions of quotients have usually been special ones devised for the particular variables under consideration, and in no way indicative of the general result. It is the purpose of this paper to study the distribution of the quotient of two variables for itself alone, with attention first to the question of existence, and then to the accurate derivation of a number of general formulas for the frequency function and d.f.[1] The principal formulas which we shall derive may be described briefly as follows (the numerals refer to the equation numbers in the text):

(3.1). The frequency function of the quotient of two variables which have an absolutely continuous joint probability function.

(4.11), (4.12). The d.f. of the quotient of a pair of arbitrary independent variables, expressed in terms of the d.f.'s of these variables.

(5.2). The d.f. of the quotient of a pair of arbitrary independent variables, expressed in terms of the c.f.'s[2] of these variables.

(6.4). The limiting form of the d.f. of a quotient of two sums of arbitrary identical independent variables.

(7.1). A formula analogous to (3.1) for the product of two chance variables.

(7.2). A formula analogous to (4.11) for the product of two chance variables.

2. **The existence of the quotient distribution.** The function $Z = X/Y$ is a continuous function of X and Y, finite and uniquely defined for all points (X, Y) such that $Y \neq 0$. Therefore if $P\{Y = 0\} = 0$, the pr.f.[3] $P(S)$ of the joint distribution of X and Y determines a probability distribution for Z (see [1, pp. 12–13]). To avoid irrelevant difficulties, we shall assume in the sequel that $P\{Y = 0\} = 0$ unless definite statement is made to the contrary. This assumption involves no real restriction on our work, for in situations in which, a priori, the assumption is not fulfilled, we can always replace the distribution

---

[1] I.e., distribution function. The underlying axioms, terminology, and abbreviations in this paper are uniform with those of Cramér's book [1]. For the definition of d.f., see [1, p. 11].

[2] I.e., characteristic functions. See [1, p. 23].

[3] I.e., probability function; [1, p. 9].

of Y by the conditional distribution of Y relative to the hypothesis that $Y \neq 0$. In such cases, then, the distribution of Z which we are about to study is to be interpreted as a conditional distribution relative to this hypothesis.

We shall suppose that the space of X is the x-axis, that of Y, the y-axis, and that of Z, the z-axis. It is quite readily seen that the set of points in the (x, y) plane which corresponds to the set $Z \leqq z$ consists of

(i) the infinite region[4] in the upper half-plane which is bounded by the negative x axis and by the line $x = zy$;

(ii) the infinite region in the lower half-plane bounded by the positive x-axis and the line $x = zy$;

(iii) the line $x = zy$ except for the origin.

Denoting this set by S_z, we have

$$H(z) = \int_{S_z} dP(S) = P(S_z),$$

where $H(z)$ is the d.f. of Z. The present paper, from the viewpoint of analysis, is simply a study of the Lebesgue-Stieltjes integral appearing in this equation.

3. The continuous case. Suppose first that $P(S)$ is absolutely continuous. This means that the joint distribution of X and Y has a frequency function $\varphi(x, y)$, which is defined almost everywhere, is non-negative, and has the property that $P(S) = \int_S \varphi(x, y)\, dx\, dy$. In general, this integral must be taken in the Lebesgue sense, but of course if the discontinuities of φ form a set of two-dimensional measure zero, and if the Jordan content of any bounded portion of the boundary of S is zero, then this integral is just an ordinary improper double Riemann integral.[5] In particular, these conditions are fulfilled if φ is continuous everywhere and if $S = S_z$.

The transformation $x = uv$, $y = v$, gives a continuous one-to-one map of S_z onto a set $\mathfrak{S}_z$ of the (u, v) plane which consists of the closed half-plane lying to the left of the line $u = z$, but with the u-axis deleted. The Jacobian of the transformation has the absolute value $|v|$. By the theorem for change of variables in Lebesgue integrals [4, pp. 653–655], we have

$$H(z) = \int_{S_z} \varphi(x, y)\, dx\, dy = \int_{\mathfrak{S}_z} |v|\, \varphi(uv, v)\, du\, dv.$$

By Fubini's Theorem [6, pp. 203–208], the last integral can be expressed as a repeated integral. Integrating first with respect to v, we obtain this result

THEOREM 3.1: *If the joint variable (X, Y) has the frequency function $\varphi(x, y)$, then*

$$H(z) = \int_{-\infty}^{z} \left[\int_{-\infty}^{+\infty} |v|\, \varphi(uv, v)\, dv \right] du,$$

---

[4] I.e., *open* connected set.

[5] See [4, pp. 476–478: p. 575].

and consequently $H(z)$ is an absolutely continuous function of z. The frequency function of the distribution of Z exists almost everywhere, and is given by the formula

$$h(z) = F'(z) = \int_{-\infty}^{+\infty} | v | \varphi(zv, v) \, dv. \tag{3.1}$$

We remark that if X and Y are independent, so that $\varphi(x, y) = f(x) \cdot g(y)$, where f and g are respectively the frequency functions of X and Y, then (3.1) may be written in the form

$$h(z) = \int_{-\infty}^{+\infty} | v | f(zv)g(v) \, dv. \tag{3.2}$$

This case was considered recently by Huntington [5], with the additional restrictions that $g(y) = 0$, $y < 0$, and that $f(x)$ and $g(y)$ be continuous.

All the familiar special quotient distributions of applied mathematical statistics, such as Student's t and Fisher's z, may conveniently and rigorously be derived by means of (3.1) and (3.2); in each case the required result follows immediately after an obvious change of variables in the integrand. We pause here only to point out explicitly the result obtained when X and Y have a normal joint distribution with variances σ_X^2, σ_Y^2, and correlation coefficient ρ. If the means $E(X)$ and $E(Y)$ are not equal to zero, it is apparently impossible to evaluate (3.1) in closed form; this case has been studied in some detail by Geary [3] and by Fieller [2]. But if $E(X) = E(Y) = 0$, then

$$h(z) = \frac{\sigma_X \sigma_Y \sqrt{1 - \rho^2}}{\pi} \cdot \frac{1}{\sigma_Y^2 \left(z - \rho \dfrac{\sigma_X}{\sigma_Y}\right)^2 + \sigma_X^2(1 - \rho^2)},$$

which is the frequency function of a Cauchy distribution with mode at the point $z = \rho\sigma_X/\sigma_Y$, the value of the regression coefficient of X on Y. If X and Y are independent, then $\rho = 0$, and the frequency function becomes

$$h(z) = \frac{\sigma_X \sigma_Y}{\pi} \cdot \frac{1}{\sigma_Y^2 z^2 + \sigma_X^2}. \tag{3.3}$$

4. **The quotient of two arbitrary independent variables.** We shall henceforth drop the restriction that $P(S)$ be absolutely continuous, but shall suppose instead that X and Y are independent chance variables with one-dimensional distributions of the most general type, except that the distribution of Y will be subject to the restriction that $P\{Y = 0\} = 0$.

We denote the d.f. of X by $F(x)$, that of Y by $G(y)$, and, as usual, that of Z

by $H(z)$. It is to be noticed that the condition $P\{Y = 0\} = 0$ implies that $G(y)$ is continuous at the point $y = 0$. Let

$$
\begin{aligned}
f(t) &= \int_{-\infty}^{+\infty} e^{itx}\, dF(x) \\
g^+(t) &= \int_0^{\infty} e^{ity}\, dG(y) \qquad (4.1)\\
g^-(t) &= \int_{-\infty}^{0} e^{ity}\, dG(y).
\end{aligned}
$$

Clearly

$$H(z) = P\{X - zY \leqq 0; Y > 0\} + P\{X - zY \geqq 0; Y < 0\}. \qquad (4.2)$$

We introduce the functions

$$
\begin{aligned}
\Gamma_1(u) &= P\{X - zY \leqq u; Y > 0\} = [1 - G(0)]\cdot P\{X - zY \leqq u \mid Y > 0\},[6] \\
\gamma_1(t) &= \int_{-\infty}^{+\infty} e^{itu}\, d\Gamma_1(u), \\
\Gamma_2(u) &= P\{zY - X \leqq u; Y < 0\} = G(0)\cdot P\{zY - X \leqq u \mid Y < 0\}, \\
\gamma_2(t) &= \int_{-\infty}^{+\infty} e^{itu}\, d\Gamma_2(u), \qquad (4.3)\\
\Gamma(u) &= \Gamma_1(u) + \Gamma_2(u) \\
\gamma(t) &= \int_{-\infty}^{+\infty} e^{itu}\, d\Gamma(u) = \gamma_1(t) + \gamma_2(t).
\end{aligned}
$$

By (4.2) and (4.3),

$$H(z) = \Gamma(0). \qquad (4.4)$$

We shall now evaluate $\Gamma_1(u)$ and $\Gamma_2(u)$ in terms of $F(x)$ and $G(y)$, and also $\gamma_1(t)$ and $\gamma_2(t)$ in terms of $f(t)$, $g^+(t)$, and $g^-(t)$.

Let us assume for a moment that $P\{Y > 0\} \neq 0$; that is, that $G(0) < 1$. The conditional distribution of Y relative to the hypothesis that $Y > 0$ then has the d.f.

$$G_1(y) = \begin{cases} \dfrac{G(y) - G(0)}{1 - G(0)}, & y \geqq 0, \\ 0, & y < 0. \end{cases} \qquad (4.5)$$

The d.f. of $-zY$ relative to this hypothesis is $G_1(-y/z)$ if $z < 0$, and $1 - G_1[(-y/z) - 0]$ if $z > 0$.

[6] By $P(A \mid b)$ is meant the conditional probability of the event A relative to the hypothesis b.

It is well known that the corresponding d.f. of the sum $X + (-zY)$ is given by a convolution of the d.f.'s of X and $(-zY)$.[7] In the present case, this result takes the form

$$(4.6)\quad P\{X - zY \leqq u \mid Y > 0\} = \begin{cases} \displaystyle\int_{-\infty}^{+\infty} F(u-v)\, dG_1\left(-\frac{v}{z}\right), & z < 0, \\ \displaystyle\int_{-\infty}^{+\infty} F(u-v)\, d\left[1 - G_1\left(-\frac{v}{z} - 0\right)\right], & z > 0. \end{cases}$$

Referring to the definition of these Lebesgue-Stieltjes integrals [4, pp. 662–663], we see that the change of variables $w = -v/z$ yields the equations

$$(4.7)\qquad P\{X - zY \leqq u \mid Y > 0\} = \begin{cases} \displaystyle\int_0^{\infty} F(u + zw)\, dG_1(w), & z < 0, \\ \displaystyle\int_0^{\infty} F(u + zw)\, dG_1(w - 0), & z > 0. \end{cases}$$

Now the definition of the variation of $G_1(y)$ [4, pp. 341–342] used in forming these Lebesgue-Stieltjes integrals makes no distinction between the variation of $G_1(y)$ and that of $G_1(y - 0)$ over any bounded set contained in an interval of integration $a \leq y < \infty$, provided that $G_1(y)$ is continuous at a in the two-sided sense. Since $G_1(y)$ is continuous at $y = 0$ in this sense, it is possible to replace $G_1(w - 0)$ by $G_1(w)$ in the second of the two integrals in (4.7).

Equation (4.7) is clearly true for $z = 0$ as well as for all other values of z. Referring to (4.5) and (4.3), we see that

$$\Gamma_1(u) = \int_0^{\infty} F(u + zw)\, dG(w), \quad \text{all } z.$$

The c.f. of the convolution (4.6) is the product of the c.f.'s of X and of the conditional distribution of $-zY$ [1, p. 36]. This product is $f(t) \cdot \int_0^{\infty} e^{-itzy}\, dG_1(y)$. Thus by (4.5), (4.3), and (4.1),

$$(4.8)\qquad \gamma_1(t) = [1 - G(0)]\left[f(t) \cdot \int_0^{\infty} e^{-itzy}\, dG_1(y)\right] = f(t)g^{+}(-tz).$$

We have established (4.7) and (4.8) under the condition that $P\{Y > 0\} \neq 0$. However, it is obvious that they are trivially true if $P\{Y > 0\} = 0$.

We turn now to $\Gamma_2(u)$. Supposing that $P\{Y < 0\} \neq 0$, the conditional distribution of Y relative to the hypothesis that $Y < 0$, has the d.f.

$$G_2(y) = \begin{cases} \dfrac{G(y)}{G(0)}, & y < 0, \\ 1, & y \geqq 0. \end{cases}$$

[7] See [1, pp. 35–36]; also [7].

The conditional distribution of zY has the d.f. $G_2(y/z)$ for $z > 0$, and $1 - G_2[(y/z) - 0]$ for $z < 0$. The d.f. of $-X$ is $1 - F(-x - 0)$. Thus

$$P\{zY - X \leqq u \mid Y < 0\} = \begin{cases} \displaystyle\int_{-\infty}^{+\infty} \{1 - F[-(u-v) - 0]\}\, d\left[1 - G_2\left(\frac{v}{z} - 0\right)\right], & z < 0, \\ \displaystyle\int_{-\infty}^{+\infty} \{1 - F[-(u-v) - 0]\}\, dG\left(\frac{v}{z}\right), & z > 0, \end{cases}$$

$$= 1 - \int_{-\infty}^{0} F(zw - u - 0)\, dG_2(w).$$

Evidently the first and last members of this equation are equal for $z = 0$ as well as for all other values of z. From (4.3) we obtain

$$\Gamma_2(u) = G(0) - \int_{-\infty}^{0} F(zw - u - 0)\, dG(w), \quad \text{all } z.$$

Also, as before,

$$\gamma_2(t) = f(-t)g^-(zt).$$

Obviously, the last two equations are still true if $P\{Y < 0\} = 0$.

To summarize, we have shown that

$$(4.9) \quad \Gamma(u) = G(0) + \int_0^{\infty} F(u + zw)\, dG(w) - \int_{-\infty}^{0} F(zw - u - 0)\, dG(w), \quad \text{all } z;$$

$$(4.10) \quad \gamma(t) = f(t)g^+(-zt) + f(-t)g^-(zt).$$

Referring now to (4.4) and letting $u = 0$ in (4.9), we are able to state the following theorem:

Theorem 4.1: *If X and Y are independent chance variables with respective d.f.'s $F(x)$ and $G(y)$, the d.f. of the quotient X/Y is given by the formula*

$$(4.11) \quad H(z) = G(0) + \int_0^{\infty} F(zw)\, dG(w) - \int_{-\infty}^{0} F(zw - 0)\, dG(w)$$

for all values of z.

We shall not attempt to make a careful study of the above formula, such as the studies which certain writers have made of convolutions. However, it does seem desirable to place on record here certain remarks concerning it of a more or less superficial character. For convenience in later reference, we state these remarks in the form of four lemmas.

Lemma 4.1: *Let M_1 be the set of all values of z such that if $z \in M_1$, the set of discontinuity points of $F(zw)$ on the w-axis has a point in common with the point spectrum of $G(w)$. Then if $z \in C(M_1)$,*[8] *the integrals $\int_0^{\infty} F(zw \pm 0)\, dG(w)$,*

[8] By $C(M_1)$ we mean the complement of M_1 with respect to the z-axis.

$\int_{-\infty}^{0} F(zw \pm 0)\, dG(w)$, *are Riemann-Stieltjes integrals and consequently the integrands can be replaced by* $F(zw)$ *without altering the values of the integrals.*

The lemma follows immediately from the definitions of Riemann-Stieltjes and Lebesgue-Stieltjes integrals.

LEMMA 4.2: *The set* M_1 *is denumerable.*

The proof can easily be supplied by the reader.

LEMMA 4.3: *Let* M_2 *be the set of all values of* z *such that if* $z \in M_2$, $\Gamma(u)$ *is discontinuous at* $u = 0$. *Then* $M_2 \subset M_1$.

To prove this statement, we first observe that $\Gamma(u)$ is a genuine d.f. [1, p. 11]. For obviously $\Gamma(-\infty) = 0$, $\Gamma(+\infty) = 1$, and since $\Gamma_1(u)$ and $\Gamma_2(u)$ are both products of d.f.'s into constants, these two functions, and therefore $\Gamma(u)$, must be continuous from the right. It is this last property of $\Gamma(u)$ which is needed for our present purposes; in particular, we have the relation $\lim_{u\to+0}\Gamma(u) = \Gamma(0) = H(z)$. On the other hand, by the general convergence theorem for Lebesgue-Stieltjes integrals [4, pp. 663–664], we have

$$\lim_{u\to-0} \Gamma(u) = G(0) + \int_{0}^{\infty} F(zw - 0)\, dG(w) - \int_{-\infty}^{0} F(zw)\, dG(w).$$

If z be chosen so that this integral and the ones in (4.11) are all Riemann-Stieltjes integrals, the expression $(zw - 0)$, wherever it appears, may be replaced by zw without changing the values of the integrals. Thus for such a value of z, $\Gamma(+0) = \Gamma(-0)$. According to Lemma 4.1, we can be sure that at least if $z \in C(M_1)$, the integrals here will be Riemann-Stieltjes integrals, so our proposition is proved.

Since $H(z_1 + 0)$ is equal to $\Gamma(+0)$ with $z = z_1$, and $H(z_1 - 0)$ is equal to $\Gamma(-0)$ with $z = z_1$, we have the following result:

LEMMA 4.4: *The set* M_2 *is the set of discontinuity points of* $H(z)$.

By using the alternate form of the convolutions used to derive (4.9), we obtain a representation of $\Gamma(u)$ somewhat more complicated than that appearing in (4.9). The corresponding formula for $H(z)$ is as follows:

$$\text{(4.12)}\quad H(z) = \begin{cases} G(0)[1 - F(-0)] - G(0)F(0) + \displaystyle\int_{-\infty}^{0} G\left(\frac{v}{z}\right) dF(v) \\ \qquad\qquad - \displaystyle\int_{0}^{\infty} G\left(\frac{v}{z} - 0\right) dF(v - 0), & z < 0; \\ F(0)[1 - G(0)] + G(0)[1 - F(-0)], & z = 0; \\ 1 + G(0)[1 - F(-0)] - G(0)F(0) + \displaystyle\int_{-\infty}^{0} G\left(\frac{v}{z}\right) dF(v - 0) \\ \qquad\qquad - \displaystyle\int_{0}^{\infty} G\left(\frac{v}{z} - 0\right) dF(v), & z > 0. \end{cases}$$

5. **Representation of $H(z)$ by characteristic functions.**. A simple algebraic formula connecting the c.f. of Z with those of X and Y is not available. However, there exists an interesting representation of $H(z)$ in terms of the functions $f(t)$, $g^+(t)$, and $g^-(t)$. The result may be stated as follows:

THEOREM 5.1:[9] *Let the distributions of the independent variables X and Y have finite first absolute moments, and let the integral*

$$\left(\int_{-\infty}^{-1} + \int_{1}^{\infty}\right) \frac{|f(t)g^+(-zt) + f(-t)g^-(zt)|}{t}\, dt \tag{5.1}$$

be finite for each value of z. Let $\Delta(u)$ be any d.f. with a finite first absolute moment, and let $\left(\int_{-\infty}^{-1} + \int_{1}^{\infty}\right) \left|\frac{\delta(t)}{t}\right| dt$ be finite, where $\delta(t)$ is the c.f. of $\Delta(u)$. Then

$$H(z) = \Delta(0) - \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \frac{f(t)g^+(-zt) + f(-t)g^-(zt) - \delta(t)}{t}\, dt. \tag{5.2}$$

If the integral obtained by formal differentiation under the integral sign with respect to z in (5.2) is uniformly convergent in a certain interval I, then the frequency function $h(z)$ of the distribution of z exists in that interval and is given by the formula

$$h(z) = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} [f(t)g^{+\prime}(-zt) - f(-t)g^{-\prime}(zt)]\, dt, \qquad z \in I.$$

We remark that the condition (5.1) will be satisfied for all values of z if $f(t)$ alone satisfies a similar condition, inasmuch as $|g^+(t)| \leqq 1$, $|g^-(t)| \leqq 1$. Important special cases of the theorem arise when $\Delta(u)$ is replaced by $F(u)$ or $G(u)$, and when $\Delta(u)$ is so chosen that $\Delta(0) = 0$.

Our proof of the theorem will depend on a rather general result due to Cramér [1, Theorem 12], which we shall restate here in the special form applicable to the problem at hand.

LEMMA 5.1: *Let $R(u)$ be a function of bounded variation over the infinite interval $-\infty < u < \infty$, let $\lim_{u\to-\infty} R(u) = \lim_{u\to+\infty} R(u) = 0$, and let $r(t) = \int_{-\infty}^{+\infty} e^{itu}\, dR(u)$. If* (a) $\int_{-\infty}^{+\infty} |u|\, dR(u)$ *and* (b) $\left(\int_{-\infty}^{-1} + \int_{1}^{\infty}\right) \left|\frac{r(t)}{t}\right| dt$, *both are finite, then for every value of u,*

$$R(u) = -\frac{1}{2\pi i} \int_{-\infty}^{+\infty} \frac{r(t)}{t} e^{-itu}\, dt.$$

To prove Theorem 5.1, we observe that since $\Gamma(u)$ is a d.f. (see proof of Lemma 4.3), the difference $\Gamma(u) - \Delta(u)$ is a function similar to the function $R(u)$ of the lemma. If we do let $R(u) = \Gamma(u) - \Delta(u)$, it follows at once that $r(t) = \gamma(t) - \delta(t) = f(t)\cdot g^+(-zt) + f(-t)g^-(zt) - \delta(t)$. If we can verify that this $R(u)$

---

[9] The theorem is due to Cramér in the case in which $G(0) = 0$, and $\Delta(u) = G(u)$. See [1, Theorem 16].

satisfies conditions (a) and (b) of the lemma, then we shall have established the relation,

$$\Gamma(u) = \Delta(u) - \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \frac{f(t)g^{+}(-zt) + f(-t)g^{-}(zt) - \delta(t)}{t} e^{-itu}\, dt,$$

for all values of u, and letting $u = 0$ in this equation, we shall obtain (5.2).

Condition (b) in the lemma is taken care of by (5.1) and the condition on $\delta(t)$ in Theorem 5.1. Clearly condition (a) will be satisfied if it turns out that $\Gamma(u)$ has a finite first absolute moment. Now the existence of finite first absolute moments of X and Y will insure the existence of finite first absolute moments for the conditional distributions involved in the definitions of $\Gamma_1(u)$ and $\Gamma_2(u)$, because $E \mid X - zY \mid \leqq E \mid X \mid + \mid z \mid E \mid Y \mid$. It follows quite readily from this that the first absolute moment of $\Gamma(u)$ is finite. The proof of the theorem is complete.

6. **Distributions of variable form.** We consider now the case in which the distributions of the numerator and denominator approach limiting forms.

THEOREM 6.1: *Let the independent variables X_α and Y_β have respective d.f.'s $F_\alpha(x)$ and $G_\beta(y)$ which depend upon the two parameters α and β. Let $H_{\alpha,\beta}(z)$ be the d.f. of the quotient $Z_{\alpha,\beta} = X_\alpha/Y_\beta$. If there exist two chance variables X and Y with respective distribution functions $F(x)$ and $G(y)$ such that $\lim_{\alpha\to\infty} F_\alpha(x) = F(x)$ at all points of continuity of $F(x)$, and $\lim_{\beta\to\infty} G_\beta(y) = G(y)$, at all points of continuity of $G(y)$, then*

$$\lim_{\substack{\alpha\to\infty\\ \beta\to\infty}} H_{\alpha,\beta}(z) = \lim_{\alpha\to\infty}\lim_{\beta\to\infty} H_{\alpha,\beta}(z) = \lim_{\beta\to\infty}\lim_{\alpha\to\infty} H_{\alpha,\beta}(z) = H(z) \tag{6.1}$$

at all points of continuity of $H(z)$, where $H(z)$ is the d.f. of the variable X/Y. The double limit in (6.1) *is uniform in any finite or infinite interval of continuity of $H(z)$.*

In the interpretation of the limits involved in this theorem, it is to be understood that in the hypotheses, α may tend to infinity over any unbounded set T_α of the α-axis, and β may tend to infinity over any unbounded set T_β of the β-axis, provided that in (6.1), α and β are restricted so that $\alpha \in T_\alpha$ and $\beta \in T_\beta$.

To prove the theorem, we introduce functions $f_\alpha(t)$, $g_\beta^{+}(t)$, $g_\beta^{-}(t)$, $\Gamma_{\alpha,\beta}(u)$, $\gamma_{\alpha,\beta}(t)$, which are defined by equations (4.1) and (4.3) with F, G, X, Y replaced respectively by F_α, G_β, X_α, Y_β. On the other hand, with reference to the distributions of X and Y, we employ the notation of section 4 without modification. According to the work in that section, $\Gamma(u)$ is given by (4.9) and its c.f. $\gamma(t)$ is given by (4.10). Also,

$$\gamma_{\alpha,\beta}(t) = f_\alpha(t)g_\beta^{+}(-zt) + f_\alpha(-t)g_\beta^{-}(zt).$$

But it is an immediate consequence of our hypotheses that $\lim_{\alpha\to\infty} f_\alpha(t) = f(t)$,

$\lim_{\beta\to\infty} g_\beta^+(t) = g^+(t)$, and $\lim_{\beta\to\infty} g_\beta^-(t) = g^-(t)$, all of the limits being uniform in any finite interval of values of t.[10] Thus

$$\lim_{\substack{\alpha\to\infty\\ \beta\to\infty}} \gamma_{\alpha,\beta}(t) = \lim_{\alpha\to\infty}\lim_{\beta\to\infty} \gamma_{\alpha,\beta}(t) = \lim_{\beta\to\infty}\lim_{\alpha\to\infty} \gamma_{\alpha,\beta}(t) = \gamma(t), \tag{6.2}$$

uniformly in any finite interval on the t-axis.

Consider the extreme members of (6.2). It follows immediately from a well-known general theorem[11] that $\lim_{\alpha\to\infty,\beta\to\infty} \Gamma_{\alpha,\beta}(u) = \Gamma(u)$ at all continuity points of $\Gamma(u)$. Then since $H_{\alpha,\beta}(z) = \Gamma_{\alpha,\beta}(0)$ and $H(z) = \Gamma(0)$, we find that

$$\lim_{\substack{\alpha\to\infty\\ \beta\to\infty}} H_{\alpha,\beta}(z) = H(z), \qquad z \in C(M_2),$$

where M_2 is the set defined in Lemma 4.3. By Lemma 4.4, the set M_2 is the set of discontinuity points of $H(z)$, so the equality of the first and last members of (6.1) is established at all continuity points of $H(z)$. The uniformity of the limit is due to a general property of convergent sequences of d.f.'s; see [1, p. 31].

The existence and equivalence to $H(z)$ of each of the iterated limits in (6.1) may be established by two consecutive applications of the foregoing argument, and by the use of (6.2). We leave the details to the reader.

It is to be remarked that both $H_{\alpha,\beta}(z)$ and $H(z)$ can be represented by (4.11), provided, of course, that F and G in (4.11) are replaced by F_α and G_β in the case of $H_{\alpha,\beta}$; thus our theorem essentially states that the order of the double limit and the integration is immaterial in this formula. A similar remark applies to formula (5.2).

The reader is reminded that we have tacitly been assuming that the d.f. of any variable appearing in a denominator is continuous at the origin. In case $G_\beta(y)$ does not satisfy this condition, but $G(y)$ does satisfy it, and if, as suggested in section 2, we consider $H_{\alpha,\beta}(y)$ to be the d.f. of the conditional distribution of $Z_{\alpha,\beta}$ relative to the hypothesis that $Y_\beta \neq 0$, then it can be shown rather easily that Theorem 6.1 remains true with this modified interpretation. But if $G(y)$ is discontinuous at the origin, and if $H(z)$ is interpreted as the d.f. of the conditional distribution, then (6.1) may be no longer true, as can be shown by trivial examples.

Perhaps the most important cases of variable distributions arise in the consideration of sums of independent chance variables. We accordingly present the following synthesis of Theorem 6.1 and a simple case of the Central Limit Theorem.

THEOREM 6.2: *Let $U_1, U_2, \cdots$, be a sequence of identically distributed chance variables, each with mean zero and (finite) standard deviation σ_U, and let V_1,*

---

[10] See [1, p. 30].

[11] See [1, Theorem 11]. The result needed here is a trivial extension of the theorem cited.

$V_2, \cdots$, be a sequence of identically distributed chance variables, each with mean zero and (finite) standard deviation σ_V. Furthermore, let the variables U_i and V_j be all independent, $i = 1, 2, \cdots, j = 1, 2, \cdots$. If m and n tend to infinity in such a way that

$$\lim_{\substack{m\to\infty \\ n\to\infty}} \sqrt{\frac{n}{m}} = k \neq 0, \tag{6.3}$$

then the d.f. of the conditional distribution of the variable

$$W_{m,n} = \frac{U_1 + U_2 + \cdots + U_m}{V_1 + V_2 + \cdots + V_n},$$

relative to the hypothesis that the denominator is different from zero, tends uniformly to the function

$$J(w) = \int_{-\infty}^{w} \frac{k\sigma_U \sigma_V}{\pi} \cdot \frac{1}{\sigma_V^2 k^2 u^2 + \sigma_U^2}\, du. \tag{6.4}$$

For if we let

$$Z_{m,n} = \frac{\dfrac{U_1 + U_2 + \cdots + U_m}{\sigma_U\sqrt{m}}}{\dfrac{V_1 + V_2 + \cdots + V_n}{\sigma_V\sqrt{n}}},$$

then $W_{m,n} = \sqrt{m/n}(\sigma_U/\sigma_V)Z_{m,n}$. The Central Limit Theorem [1, Theorem 20] states that the d.f.'s of the numerator and denominator of $Z_{m,n}$ each tend to the function $\int_{-\infty}^{x} (1/\sqrt{2\pi})e^{-t^2/2}\, dt$, which is the d.f. of a normal distribution with mean zero and variance one. By (3.3), the quotient of two variables, each of which has this d.f., has the continuous d.f. $H(z) = \int_{-\infty}^{z} (1/\pi)[1/(1 + x^2)]\, dx$. If we let $H_{m,n}(z)$ denote the d.f. of the conditional distribution of $Z_{m,n}$, relative to the hypothesis that the denominator of $Z_{m,n}$ is different from zero, then by Theorem 6.1, $\lim_{m\to\infty, n\to\infty} H_{m,n}(z) = H(z)$ uniformly in z. Now the d.f. of the conditional distribution of $W_{m,n}$ is $H_{m,n}[\sqrt{n/m}(\sigma_V/\sigma_U)w]$, and because of (6.3) and the uniformity of the limit of $H_{m,n}(z)$, this approaches $H[k(\sigma_V/\sigma_U)w]$. Differentiating the last expression with respect to w, we find that the resulting frequency function is equal to $J'(w)$; and this concludes the proof.

As an application of the theorem, let us consider the following problem. From an urn containing white and black balls in the proportion of p to $1 - p$, we shall make 100 random drawings of a single ball with replacement after each drawing. Let $W_{50,50}$ be the ratio of the deviation of the number of white balls in the first 50 drawings from the expected number, to the deviation of the number of white balls in the second 50 drawings from the expected number. What is

the approximate value of w for which $P\{W_{50,50} \geqq w \mid b\} = .05$, where the hypothesis b is that the denominator of $W_{50,50}$ shall be different from zero?[12]

To answer this question, we observe that the numerator and denominator of $W_{50,50}$ can each be expressed as the sum of 50 independent identical chance variables, each with mean zero and with variance $p(1 - p)$. Thus according to Theorem 6.2, the approximate d.f. of $W_{50,50}$ is

$$J(w) = \int_{-\infty}^{w} \frac{1}{\pi} \frac{1}{1 + u^2} du = \frac{1}{2} + \frac{1}{\pi} \arctan w,$$

and the required value of w satisfies the equation $J(\infty) - J(w) = .05$. The solution of this equation (correct to one decimal place) is $w = 6.3$.

It is perhaps needless to remark that a study of the error involved in supposing $J(w)$ to be the d.f. of $W_{m,n}$ in Theorem 6.2, must necessarily precede the unreserved acceptance of numerical results obtained by means of that theorem.

7. **Products of chance variables.** We conclude this paper with a rather brief treatment of the distribution of the product of two chance variables. To preserve a notation uniform with that of the preceding sections, we shall write the product as $X = YZ$, where the d.f.'s of X, Y, and Z are to be denoted, as before, by $F(x)$, $G(y)$, and $H(z)$, respectively. The existence of $F(x)$ is readily proved by the methods of section 2. The assumption that $P\{Y = 0\} = 0$ is of course unnecessary here, and will be dropped in this section.

In the continuous case, an argument similar to the one employed in section 3 will establish the following result:

THEOREM 7.1: *If the joint variable (Y, Z) has the frequency function $\psi(y, z)$, then*

$$F(x) = \int_{-\infty}^{x} \left[\int_{-\infty}^{+\infty} \left| \frac{1}{v} \right| \psi\left(\frac{u}{v}, v\right) dv \right] du$$
$$= \int_{-\infty}^{x} \left[\int_{-\infty}^{+\infty} \left| \frac{1}{v} \right| \psi\left(v, \frac{u}{v}\right) dv \right] du,$$

and consequently $F(x)$ is an absolutely continuous function of x. The frequency function of the distribution of X exists almost everywhere, and is given by the formula

$$f(x) = F'(x) = \int_{-\infty}^{+\infty} \left| \frac{1}{v} \right| \psi\left(\frac{x}{v}, v\right) dv = \int_{-\infty}^{+\infty} \left| \frac{1}{v} \right| \psi\left(v, \frac{x}{v}\right) dv. \tag{7.1}$$

In the discontinuous case, with Y and Z independent, we can write $X = ZY = Z/(1/Y)$ and use Theorem 4.1 to derive a formula for $F(x)$. We have:

$$F(x) = P\{X \leqq x\} = P\{Y \neq 0\}P\{X \leqq x \mid Y \neq 0\} + P\{X \leqq x; Y = 0\}.$$

[12] This hypothesis would always be fulfilled in case $50p$ is not an integer.

Excluding for a moment the trivial case in which $P\{Y \neq 0\} = 0$, let $G_1(y)$ be the d.f. of the conditional distribution of $(1/Y)$ relative to the hypothesis that $Y \neq 0$. Then

$$P\{Y \neq 0\}G_1(y) = \begin{cases} G(-0) + 1 - G\left(\frac{1}{y} - 0\right), & y > 0, \\ G(-0), & y = 0, \\ G(-0) - G\left(\frac{1}{y} - 0\right), & y < 0. \end{cases}$$

It is to be observed that $G_1(y)$ is continuous at $y = 0$. Using Theorem 4.1, we find that

$$P\{X \leqq x \mid Y \neq 0\} = G_1(0) + \int_0^\infty H(xw)\, dG_1(w) - \int_{-\infty}^0 H(xw - 0)\, dG_1(w).$$

So

$$\begin{aligned} &P\{Y \neq 0\}P\{X \leqq x \mid Y \neq 0\} \\ &= G(-0) + \int_{0+0}^\infty H(xw)\, d\left[-G\left(\frac{1}{w} - 0\right)\right] - \int_{-\infty}^{0-0} H(xw - 0)\, d\left[-G\left(\frac{1}{w} - 0\right)\right] \\ &= G(-0) + \int_{0+0}^\infty H\left(\frac{x}{v}\right) dG(v) - \int_{-\infty}^{0-0} H\left(\frac{x}{v} - 0\right) dG(v). \end{aligned}$$

This equation is trivially true if $P\{Y \neq 0\} = 0$. Also,

$$P\{X \leqq x;\, Y = 0\} = \begin{cases} 0, & x < 0, \\ G(0) - G(-0), & x \geqq 0. \end{cases}$$

Thus we obtain the following theorem:

THEOREM 7.2: *If Y and Z are independent chance variables with respective d.f.'s $G(y)$ and $H(z)$, then the d.f. of their product is given by the formula*

$$F(x) = \int_{0+0}^\infty H\left(\frac{x}{v}\right) dG(v) - \int_{-\infty}^{0-0} H\left(\frac{x}{v} - 0\right) dG(v) + \begin{cases} G(-0), & x < 0, \\ G(0), & x \geqq 0, \end{cases} \tag{7.2}$$

for all values of x.

REFERENCES

[1] H. CRAMÉR, *Random Variables and Probability Distributions*, Cambridge, 1937.

[2] E. C. FIELLER, "The distribution of the index in a normal bivariate population," *Biometrika*, Vol. 24 (1932), pp. 428–440.

[3] R. C. GEARY, "The frequency distribution of the quotient of two normal variates," *Roy. Stat. Soc. Jour.*, Vol. 93 (1930), pp. 442–446.

[4] E. W. HOBSON, *The Theory of Functions of a Real Variable*, Vol. 1, Cambridge, 1927.

[5] E. V. HUNTINGTON, "Frequency distribution of product and quotient," *Annals of Math. Stat.*, Vol. 10 (1939), pp. 195–198.

[6] H. KESTELMAN, *Modern Theories of Integration*, Oxford, 1937.

[7] A. WINTNER, "On the addition of independent distributions," *Am. Jour. Math.*, Vol. 56 (1934), pp. 8–16.

SOME GENERALIZATIONS OF THE LOGARITHMIC MEAN AND OF SIMILAR MEANS OF TWO VARIATES WHICH BECOME INDETERMINATE WHEN THE TWO VARIATES ARE EQUAL

By Edward L. Dodd

University of Texas

1. Introduction. The logarithmic mean m of positive numbers, x and y, as given by

$$m = \frac{y - x}{\log_e y - \log_e x} = \frac{y - x}{\log_e (y/x)} \tag{1}$$

is of considerable importance in problems[1] relating to the flow of heat.

The logarithmic mean arises, moreover, in less technical problems such as the following: Given that incomes t in the interval, $x \leqq t \leqq y$, are distributed with frequency inversely proportional to t. That is, with k = a positive constant,

$$\phi(t)\, dt = (k/t)\, dt \tag{2}$$

is the number of individuals with incomes lying between t and $t + dt$. Then, with $x > 0$, the total number f of individual incomes is

$$f = \int_x^y \phi(t)\, dt = k(\log y - \log x). \tag{3}$$

The combined income g of the group is

$$g = \int_x^y t\phi(t)\, dt = k(y - x). \tag{4}$$

And thus the *logarithmic* mean g/f of the *two* numbers x and y in (1) is the *arithmetic mean of all* the incomes; that is, the *average income*—at least to a close approximation if the group is large enough that integration may replace summation.

Now m in (1) becomes *indeterminate*, if $x = y$. Nevertheless, if $c > 0$, and $x \to c$ and $y \to c$, then $m \to c$. Thus, we may properly speak of m as a mean of these two variates, x and y.

This logarithmic mean is one of a set of means studied by Renzo Cisbani[2], the general form being

[1] See Walker, Lewis, and McAdams, *Principles of Chemical Engineering*, McGraw Hill & Co., Part IV, Logarithmic mean temperature difference.

[2] R. Cisbani, "Contributi alla teoria delle medie." *Metron*, Vol. 13(1938), pp. 23-34.

$$z = \left[\frac{b^{x+j} - a^{x+j}}{(x/j + 1)(b^j - a^j)}\right]^{1/x} \tag{5}$$

and the logarithmic mean appearing when $x = 1, j \to 0$.

In a chart between pages 28 and 29 Cisbani exhibits thirty varieties of these means (5). It will be noticed that z is *indeterminate* if $a = b$.

Some methods for dealing with means which may become indeterminate forms I have indicated in a recent paper.[3]

Now a generalization from a mean of *two* variates to a mean of three or more variates may sometimes *seem* to be *immediate*. However, for the arithmetic mean $(x + y)/2$ of two variates x and y, the function [min. (x, y, z) + max. $(x, y, z)]/2$ is as much a generalization as is the arithmetic mean $(x + y + z)/3$. *Actually*, the direction in which generalization is to take place is *arbitrary*. However, it is natural to expect the generalization to arise from a problem somewhat similar to one that may give rise to the original mean. And it is desirable that to the generalization should be carried over as many properties or characteristics of the original as is possible.

In the foregoing illustration, we considered a *single* interval $x \leqq t \leqq y$ in which incomes are distributed in accordance with a *relative* frequency proportional to $\phi(t)$. And the *arithmetic* mean of *all* these incomes was obtained as a *logarithmic* mean of the *two* range limits x and y, at least approximately, allowing integration to take the place of summation. If $\phi(t)$ had been $kt^{-3/2}$, instead of kt^{-1}, then the average of *all* the incomes would have been the *geometric* mean of the *two* range limits x and y.

To effect a *first* generalization, we shall now suppose an original interval x_0 to x_n, to be divided into n subintervals by points x_r such that

$$x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n . \tag{6}$$

For each subinterval x_{r-1} to x_r the *same* function $\phi(t)$ will be used to describe the *relative* frequency; but the *total population* for this subinterval will be controlled by a positive constant k_r, in general *different* for the different subintervals. This may be described as *stratification*. To make this more concrete, let us suppose, as before, that $\phi(t) = k/t$. Then, with $x_0 > 0$, the mean M, which will be described more in detail in the next section, will take the form

$$M = \frac{\sum_1^n k_r(x_r - x_{r-1})}{\sum_1^n k_r \log (x_r/x_{r-1})} . \tag{7}$$

Applied to incomes, M would, like m in (1), give average income. To get some idea of the significance of k_r, let us imagine that in some community there are f_r individuals in the income bracket x_{r-1} to x_r, say from \$1001 to \$2000. Let us suppose now that f_r other individuals with incomes between \$1001 and \$2000 distributed in exactly the same manner move into this same community.

[3] "The substitutive mean and certain subclasses of this general mean." *Annals of Math. Stat.*, Vol. 11(1940), pp. 163–176. See p. 171.

Then k_r would be changed to $k_r' = 2k_r$. But, of course, among the entire $2f_r$ individuals the *relative* distribution of incomes is *exactly* the *same* as among the original f_r individuals.

In *this interpretation* k_r is a *weight* for a *bracket of items*. But, taking M in (7) just as it stands, k_r is the *weight* for the *consecutive pair* of *numbers* x_{r-1} and x_r.

2. The first generalization. When t is in some interval, $I = (a, a')$, finite or infinite, let $\phi(t)$ be a non-negative, integrable function of t.

And in I let the points at which $\phi(t) = 0$, if any, form a null-set. Then, with t in I, write

$$\Phi(t) = \int_a^t \phi(t)\,dt. \tag{8}$$

And, supposing that in (6), $a < x_0$, $a_n < a'$, set

$$f_r = \int_{x_{r-1}}^{x_r} \phi(t)\,dt = \Phi(x_r) - \Phi(x_{r-1}); \qquad r = 1, 2, \cdots, n. \tag{9}$$

Then $f_r > 0$; since $\phi(t) > 0$ and is continuous almost everywhere in (x_{r-1}, x_r). Since in any finite subinterval of I, $t\phi(t)$ is integrable, we may set

$$\Psi(t) = \int_a^t \psi(t)\,dt = \int_a^t t\phi(t)\,dt. \tag{10}$$

$$g_r = \int_{x_{r-1}}^{x_r} \psi(t)\,dt = \Psi(x_r) - \Psi(x_{r-1}). \tag{11}$$

Now, by a mean value theorem, there exists a number t_r' such that

$$g_r/f_r = t_r', \qquad x_{r-1} < t_r' < x_r. \tag{12}$$

Taking *positive* numbers k_r, the weighted arithmetic mean of g_r/f_r, with weights k_rf_r is then

$$M = \frac{\sum_1^n k_r g_r}{\sum_1^n k_r f_r} = \frac{\sum_1^n k_r[\Psi(x_r) - \Psi(x_{r-1})]}{\sum_1^n k_r[\Phi(x_r) - \Phi(x_{r-1})]}. \tag{13}$$

If $\phi(t) = k/t$, this becomes the mean (7) associated with the logarithmic mean. Now, since for (13) the weights k_rf_r are *positive*, it follows from (12) that

$$x_0 < t_1' \leqq M \leqq t_n' < x_n. \tag{14}$$

Suppose, now, that b lies in I, and that subject to (6) each $x_r \to b$. Then, by (14), $M \to b$. And thus M is an *internal mean* of $x_0, x_1, \cdots, x_n$, although with the x's all equal, M assumes an indeterminate form.

In (13) the *weights* k_r are applied to *pairs* of numbers, either to $\Psi(x_r) - \Psi(x_{r-1})$ or to $\Phi(x_r) - \Phi(x_{r-1})$, whereas in *most* weighted means, the weights are applied

to *individual* numbers. We consider now a form equivalent to (13), but in which the weights c_r are attached to the *individual* numbers. It seemed possible to get a more general mean than (13) by abandoning certain conditions upon the weights c_r which first arose. But such relaxing of restrictions leads to difficulties, as will be shown. By setting

$$(15) \qquad c_0 = -k_1, \quad c_n = k_n; \quad c_r = k_r - k_{r+1}, \quad r = 1, 2, \cdots, n-1,$$

we may write M in the form;

$$(16) \qquad M = \frac{\sum_0^n c_r \Psi(x_r)}{\sum_0^n c_r \Phi(x_r)}.$$

On the other hand, if we choose c's subject to

$$(17) \qquad c_0 < 0, \quad c_r < -(c_0 + c_1 + \cdots + c_{r-1}) \quad \text{for } 0 < r < n,$$

$$(18) \qquad c_n = -\sum_0^{n-1} c_r;$$

then positive k's can be found to pass from (16) back to (13).

The question arises whether if the conditions (17) are abandoned, and with the c_r not all zero, (18) is retained as

$$(19) \qquad \sum_0^n c_r = 0; \qquad \text{Some } c_r \neq 0,$$

M in (16) will continue to be a mean of $x_0, x_1, \cdots, x_n$, possibly, an external mean.

It may be noted that the condition $\sum c_r = 0$ arises from the fact that when parentheses are removed from (13), each k_r is matched by $-k_r$.

By an example, it will be shown that under (19) alone, M in (16) may fail to be a mean. In (8) and (10) take $a = 0$. Then with $n = 2$, $\phi(t) = t$, take $c_0 = 1$, $c_1 = -2$, $c_2 = 1$ in (16). Then

$$(20) \qquad M = \frac{x_0^2 - 2x_1^2 + x_2^2}{2(x_0 - 2x_1 + x_2)}.$$

If $b > 0$, $\epsilon = x_0 - b$, $\eta = x_1 - b$, and $\xi = x_2 - b$, then

$$(21) \qquad M = b + \frac{1}{2}\,\frac{\epsilon^2 - 2\eta^2 + \xi^2}{\epsilon - 2\eta + \xi}.$$

If now $\eta = 2\epsilon$, and $\xi = 3\epsilon + \epsilon^2$, then

$$(22) \qquad M = b + (2 + 6\epsilon + \epsilon^2)/2 \to b + 1, \qquad \text{as } \epsilon \to 0.$$

Since M does not approach b here, when x_0, x_1, and $x_2 \to b$, in the manner specified, M in (20) is *not* a mean of x_0, x_1, and x_2.

We may enquire, further, whether the function M in (16) could be a mean if, discarding (13), (17) and (18), we put upon c_r the single restriction $c_r > 0$. In that case, if $x_0 < t < x_n$, then, since $\Phi(t)$ and $\Psi(t)$ are continuous functions of t—see (8), (10)—it would follow that if each $x_r \to t$, then $M \to \Psi(t)/\Phi(t)$. But

if M is to be a mean of $x_0, x_1, \cdots, x_n$, then $M \to t$ when each $x_r \to t$. Thus we are led to $\Psi(t) = t\Phi(t)$. Except possibly for points of a null set, $\Phi(t)$ and $\Psi(t)$ have derivatives $\phi(t)$ and $\psi(t)$; and thus

$$(23) \qquad \psi(t) = \Psi'(t) = t\Phi'(t) + \Phi(t) = t\phi(t) + \Phi(t).$$

But then, since $\psi(t) = t\phi(t)$—see (10)—it would follow that $\Phi(t) = 0$ almost everywhere in I; but $\Phi(t) > 0$, if $t > a$. Hence the assumption $c_r > 0$ is not sufficient to make the function in (16) a mean of $x_0, x_1, \cdots, x_n$.

In the simple case of $n = 1$, M becomes

$$(24) \qquad M = \frac{\Psi(x_1) - \Psi(x_0)}{\Phi(x_1) - \Phi(x_0)};$$

and this is a *symmetrical* function of x_0 and x_1.

The question arises whether if $n > 1$, M in (13) or (16) can be a symmetrical function of $x_0, x_1, \cdots, x_n$. Assume, if possible, that with $x < y < z$,

$$(25) \qquad H(x, y, z) \equiv \frac{c_0\Psi(x) + c_1\Psi(y) + c_2\Psi(z)}{c_0\Phi(x) + c_1\Phi(y) + c_2\Phi(z)}$$

is a symmetrical function of x, y and z. Now if $a/b = c/d$, and $b - d \neq 0$, it is well known that $a/b = (a - c)/(b - d)$.

Hence, if $H(x, y, z) = H(z, y, x)$, and $c_0 \neq c_2$, then

$$(26) \qquad H(x, y, z) \equiv \frac{(c_0 - c_2)\,[\Psi(x) - \Psi(z)]}{(c_0 - c_2)\,[\Phi(x) - \Phi(z)]},$$

which is not symmetrical in the three variables. Then H is not symmetrical in x, y and z, unless, possibly, when $c_0 = c_2$.

Likewise from $H(x, y, z) = H(x, z, y)$, we are led to the conclusion that H is *not* a symmetrical function of x, y, and z, unless possibly when $c_1 = c_2$. But $c_0 = c_1 = c_2$ substituted into (15) makes $k_1 = k_2 = 0$, which is contrary to hypothesis that $k_r > 0$. Then in (25) the constants c_0, c_1 and c_2 can not be chosen in conformity with (15) so as to make $H(x, y, z)$ a symmetrical function of the *three* variables.

Symmetry in *two* variables will appear, however, if the mean (13) *reduces* to a mean of just *two* variables as it does when each $k_r = k$, constant, in which case,

$$(27) \qquad M = \frac{\Psi(x_n) - \Psi(x_0)}{\Phi(x_n) - \Phi(x_0)}.$$

Although in the generalization (13) *symmetry* is thus lost, another property, *homogeneity* is retained in what seem to be the most important cases.

Most means $\Omega(x, y, \cdots, w)$ in common use are *homogeneous* functions of their arguments. That is, if c is a constant, and $\Omega(x, y, \cdots, w)$ and $\Omega(cx, cy, \cdots, cw)$ are both defined when $x, y, \cdots, w$ lie in some interval J, then

$$\Omega(cx, cy, \cdots, cw) = c\Omega(x, y, \cdots, w). \tag{28}$$

This *homogeneity* is associated geometrically with ruled surfaces, in particular with *cones.*

With reference to (8) and (10), let us write

$$F(x, y) = \frac{\Psi(y) - \Psi(x)}{\Phi(y) - \Phi(x)}. \tag{29}$$

And now, let us consider a special variety of means obtained by taking in (8)

$$\phi(t) = t^q, \tag{30}$$

where q is any real number. Then $F(x, y)$ is a homogeneous mean; that is,

$$F(cx, cy) = cF(x, y). \tag{31}$$

This is valid, indeed, even in the special cases, $q = 0, -1$, and -2, which lead, respectively to the arithmetic mean, the logarithmic mean (1) and to a second variety of logarithmic mean

$$m = \frac{xy \log (y/x)}{y - x}, \tag{32}$$

exhibited by Cisbani. It may be noted that $q = -3/2$ leads to the geometric mean, and $q = -3$ to the harmonic mean of x and y.

It is conceivable that for $\phi(t)$ other functions than t^q—functions not equivalent to t^q in integration—might be used to lead to a homogeneous $F(x, y)$ in (29). But such functions, if any, would hardly seem to be in common use.

The M in (13) retains the property of homogeneity, at least for $\phi(t) = t^q$; and so will also the more general means exhibited in the next section.

3. Further generalization. The means of Cisbani (5) suggest the following generalization. Let p be an integer or the reciprocal of an odd integer. With the notation of (13), take $k_r > 0$, and

$$F_p = \textstyle\sum_1^n k_r f_r^p, \qquad G_p = \sum_1^n k_r g_r^p, \tag{33}$$

$$M_p = [G_p/F_p]^{1/p}. \tag{34}$$

Indeed, if in (8) and (10), $a \geqq 0$, then $g_r > 0$; and we may take for p any real number except zero. Now, M_p^p may be described as the weighted arithmetic mean of $(g_r/f_r)^p$ with *positive* weights $k_r f_r^p$. And hence M_p is an *internal* mean of $x_0, x_1, \cdots, x_n$; that is

$$x_0 \leqq M_p \leqq x_n. \tag{35}$$

Furthermore, if in (8), $\phi(t) = t^q$, where q is any real number, then M_p is a homogeneous mean of $x_0, x_1, \cdots, x_n$.

Another generalization may be obtained by writing

$$m_r = g_r/f_r, \tag{36}$$

$$M'_p = [\Sigma k_r m_r^p / \Sigma k_r]^{1/p}. \tag{37}$$

And still another

$$M'' = [m_1^{k_1} \cdot m_2^{k_2} \cdots m_n^{k_n}]^{1/\Sigma k_r}. \tag{38}$$

These means (37) and (38) are internal; and they are homogeneous, if $F(x, y)$ in (29) is homogeneous.

The foregoing means are not, for $n > 1$, symmetrical functions of $x_1, x_2, \cdots, x_n$. Now the mere abandonment of (6) may lead to functions like (20) which are not means at all. But symmetry may be introduced as follows. First, lay aside (6), but suppose that the x_r are all different. Then let

$$f_{r,s} = \int_{x_r}^{x_s} \phi(t)\, dt, \qquad g_{r,s} = \int_{x_r}^{x_s} t\phi(t)\, dt; \tag{39}$$

where $r = 0, 1, \cdots, (n - 1)$; $r < s \leqq n$. Then, let

$$U = \Sigma f_{r,s}^2, \qquad V = \Sigma g_{r,s}^2; \tag{40}$$

where U and V is each a sum of $n(n - 1)/2$ terms: Let W be the double-valued mean

$$W = \pm[V/U]^{1/2}. \tag{41}$$

Then W is a symmetric function of $x_0, x_1, \cdots, x_n$. If, in (8), $a' \leqq 0$, then in (12) each $g_r/f_r < 0$; and in (41) the negative value of W is an internal mean. But the positive radical is *external*. On the other hand, if $a \geqq 0$; then $g_r/f_r > 0$; and the positive radical in (41) is internal. In this case, it may be well to use for W *only the positive value of* W.

In the more general case where $a < 0$ and $a' > 0$, the fractions g_r/f_r may have different signs. But, in all cases, *at least one* of the two radicals (41) is an *internal* mean of $x_0, x_1, \cdots, x_n$. Moreover, W is homogeneous, if in (8), $\phi(t) = t^q$.

Finally, let

$$m_{r,s} = g_{r,s}/f_{r,s}, \tag{42}$$

$$Z = \pm\{[\Sigma m_{r,s}^2]/n(n - 1)\}^{1/2}. \tag{43}$$

Then Z is symmetric; and at least one value is internal. If $a > 0$, we would naturally take $Z > 0$; and this Z is then an internal mean. Moreover, Z is homogeneous if the $m_{r,s}$ are homogeneous; that is, if $F(x, y)$ in (29) is homogeneous for every x and y in I.

A STUDY OF R. A. FISHER'S z DISTRIBUTION AND THE RELATED F DISTRIBUTION[1]

By Leo A. Aroian

Hunter College

1. Nature of the problem. Consider two samples of N_1 and N_2 drawings, each sample drawn from one of two populations consisting of variates normally distributed with equal population variances σ^2. We define the two sample means $\bar{x}_1 = \frac{\sum_{i=1}^{N_1} x_i}{N_1}$, $\bar{x}_2 = \frac{\sum_{j=1}^{N_2} x_j}{N_2}$, x_i's and x_j's independent variates. We calculate from the two samples

$$s_1^2 = \frac{\sum_{i=1}^{N_1} (x_i - \bar{x}_1)^2}{n_1} \quad \text{and} \quad s_2^2 = \frac{\sum_{j=1}^{N_2} (x_j - \bar{x}_2)^2}{n_2}, \qquad n_1 = N_1 - 1,\ n_2 = N_2 - 1.$$

The distribution of $z = \frac{1}{2} \log \frac{s_1^2}{s_2^2}$ is well known.

$$P(z) = \frac{2 n_1^{\frac{1}{2} n_1} n_2^{\frac{1}{2} n_2}}{B\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \frac{e^{n_1 z}}{(n_1 e^{2z} + n_2)^{\frac{1}{2}(n_1 + n_2)}} dz. \tag{1.1}$$

We shall denote the ordinates by $y(z)$. The purpose of this study is to discuss the seminvariants of the z distribution and also to find useful approximations for them; to show that as n_1 and n_2 approach infinity in any manner whatever the distribution of z approaches normality; to find the upper bound of the absolute value of the difference between the distribution function of z and the function determined by the approximate seminvariants of the distribution of z for n_1 and n_2 large; to approximate the z distribution by the Type III distribution, the Gram-Charlier Type A series, and the logarithmic frequency curve; and finally to investigate the same properties with respect to the F distribution, where $F = e^{2z} = \frac{s_1^2}{s_2^2}$. The non-existence of the moments of F for certain values of n_1 and n_2 is noted and explained on the basis of the distribution of the quotient $\frac{y}{x}$.

[1] Presented to the American Mathematical Society, September 10, 1938, New York City in part; and to the Institute December 27, 1939 at Philadelphia.

2. General features of the z distribution. The z distribution is always unimodal, asymmetrical if $n_1 \neq n_2$, and symmetrical if $n_1 = n_2$. We see that interchanging n_1 and n_2 is the same as replacing z by $-z$. Fisher [7] noted that the two parameter family of curves includes as special cases the normal curve, the χ^2 distribution, and Student's distribution. The mode is at $z = 0$, the maximum ordinate is

$$y(0) = \frac{2n_1^{\frac{1}{2}n_1} n_2^{\frac{1}{2}n_2}}{\mathrm{B}\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} (n_1 + n_2)^{-\frac{1}{2}(n_1+n_2)}$$

or approximately

$$(2.1) \qquad y(0) = \frac{1}{\sqrt{2\pi}} \left\{\frac{1}{2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right\}^{-\frac{1}{2}} \quad \text{for } n_1 \text{ and } n_2 \text{ large.}$$

The two points of inflection are

$$(2.2) \quad z = \tfrac{1}{2} \log \left\{\frac{n_1 n_2 + n_1 + n_2 \pm \sqrt{n_1^2 + n_2^2 + 2n_1^2 n_2 + 2n_1 n_2^2 + 2n_1 n_2}}{n_1 n_2}\right\}.$$

They are equidistant from the mode, a property also of the Pearson system of frequency curves [24]. Also $\lim\limits_{z \to \pm\infty} z^n \dfrac{d^n y(z)}{dz^n} = 0$.

3. The moment generating function and seminvariants. The moment generating function of the z distribution is

$$(3.1) \quad M_z(\theta) = \left(\frac{n_2}{n_1}\right)^{\frac{1}{2}\theta} \frac{\mathrm{B}\left(\frac{n_2 - \theta}{2}, \frac{n_1 + \theta}{2}\right)}{\mathrm{B}\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} = \left(\frac{n_2}{n_1}\right)^{\frac{1}{2}\theta} \frac{\Gamma\left(\frac{n_2 - \theta}{2}\right)\Gamma\left(\frac{n_1 + \theta}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}.$$

The seminvariants of Thiele are defined by the following identity in θ:

$$(3.2) \qquad \log M_z(\theta) = \lambda_1\theta + \lambda_2 \frac{\theta^2}{2!} + \lambda_3 \frac{\theta^3}{3!} + \lambda_4 \frac{\theta^4}{4!} + \cdots.$$

To find λ_r we take the logarithm of the moment generating function, expand it in powers of θ and choose the coefficient of $\dfrac{\theta^r}{r!}$. A complete discussion of properties of seminvariants may be found elsewhere [4].

4. The seminvariants of z. Now by the following formulas [11] p. 38:

$$(4.1) \quad \log \Gamma(1 + x) = \frac{-s_1 x}{1} + \frac{s_2 x^2}{2} - \frac{s_3 x^3}{3} + \frac{s_4 x^4}{4} - \cdots, \qquad |x| < 1,$$

$$\log \Gamma(1 - x) = s_1 x + \frac{s_2 x^2}{2} + \frac{s_3 x^3}{3} + \frac{s_4 x^4}{4} + \cdots, \qquad |x| < 1, \tag{4.2}$$

where in both formulas

$$s_1 = \lim_{n\to\infty}\left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n} - \log n\right),$$

$$s_n = \frac{1}{1^n} + \frac{1}{2^n} + \frac{1}{3^n} + \frac{1}{4^n} + \cdots, \qquad n \geqq 2.$$

Also

$$\log \mathrm{B}(\tfrac{1}{2}[1 + x], \tfrac{1}{2}) = \log \pi - \sigma_1 x + \sigma_2 \frac{x^2}{2} - \sigma_3 \frac{x^3}{3} + \sigma_4 \frac{x^4}{4} - \cdots, \qquad |x| < 1, \tag{4.3}$$

where

$$\sigma_n = \frac{1}{1^n} - \frac{1}{2^n} + \frac{1}{3^n} - \frac{1}{4^n} + \cdots, \qquad n \geqq 1$$

and

$$\sigma_n = \left(1 - \frac{1}{2^{n-1}}\right) s_n, \qquad n \geqq 2.$$

Hence from (4.1) and (4.3)

$$\log \Gamma\left(\frac{1 + x}{2}\right) = \tfrac{1}{2}\log \pi - x\left(\sigma_1 + \frac{s_1}{2}\right) + \frac{x^2}{2}\left(\sigma_2 + \frac{s_2}{2^2}\right) - \frac{x^3}{3}\left(\sigma_3 + \frac{s_3}{2^3}\right) + \frac{x^4}{4}\left(\sigma_4 + \frac{s_4}{2^4}\right) - \cdots. \tag{4.4}$$

Since $\sigma_n = \left(1 - \frac{1}{2^{n-1}}\right) s_n$, $n \geqq 2$, we may write (4.4) as

$$\log \Gamma\left(\frac{1 + x}{2}\right) = \tfrac{1}{2}\log \pi - x\left(\sigma_1 + \frac{s_1}{2}\right) + \sum_{k=2}^{\infty} \frac{(-1)^k x^k}{k}\left(1 - \frac{1}{2^k}\right) s_k. \tag{4.5}$$

From (3.1)

$$\log M_z(\theta) = \log \Gamma\left(\frac{n_2 - \theta}{2}\right) + \log \Gamma\left(\frac{n_1 + \theta}{2}\right) + \frac{\theta}{2}(\log n_2 - \log n_1) - \log \Gamma\left(\frac{n_1}{2}\right) - \log \Gamma\left(\frac{n_2}{2}\right). \tag{4.6}$$

The results assume slightly different forms for (A) n_1 and n_2 each even; (B) n_1 and n_2 each odd; (C) n_1 even, n_2 odd; (D) n_1 odd, n_2 even. The general formula for $\lambda_{r:z}$ for all cases is

$$\lambda_{r:z} = \sum_{k=0}^{\infty} \left\{ \frac{(-1)^r (r-1)!}{(n_1+2k)^r} + \frac{(r-1)!}{(n_2+2k)^r} \right\}, \qquad r \geqq 2. \tag{4.7}$$

This result is not so useful from the point of view of numerical applications as the formulas which follow.

5. Case A, n_1 and n_2 each even. From (4.6)

$$\begin{aligned} \log \Gamma\left(\frac{n_2-\theta}{2}\right) &= \log\left(\frac{n_2-2-\theta}{2}\right) + \log\left(\frac{n_2-4-\theta}{2}\right) + \cdots \\ &\quad + \log\left(1-\frac{\theta}{2}\right) + \log\Gamma\left(1-\frac{\theta}{2}\right). \end{aligned} \tag{5.1}$$

Now $\log\left(1 - \dfrac{\theta}{n_2-2}\right) = -\sum_{k=1}^{\infty} \dfrac{1}{k}\left(\dfrac{\theta}{n_2-2}\right)^k$. There will be $\dfrac{n_2}{2} - 1$ series of this sort, and only one series of the type $\log\Gamma\left(1-\dfrac{\theta}{2}\right) = \sum_{k=1}^{\infty} \dfrac{s_k}{k}\left(\dfrac{\theta}{2}\right)^k$ as given by (4.1). In the above expansion and those succeeding, terms not involving θ are omitted, since such terms are not needed in finding the seminvariants of z. The series $\log\Gamma\left(1-\dfrac{\theta}{2}\right)$ will always occur. Then

$$\begin{aligned} \log\Gamma\left(\frac{n_2-\theta}{2}\right) &= -\sum_{k=1}^{\infty}\frac{1}{k}\left[\left(\frac{\theta}{n_2-2}\right)^k + \left(\frac{\theta}{n_2-4}\right)^k + \cdots \right. \\ &\quad \left. + \left(\frac{\theta}{2}\right)^k - s_k\left(\frac{\theta}{2}\right)^k\right], \end{aligned} \tag{5.2}$$

or

$$\log\Gamma\left(\frac{n_2-\theta}{2}\right) = \sum_{k=1}^{\infty}\frac{s_k}{k}\left(\frac{\theta}{2}\right)^k - \sum_{k=1}^{\infty}\frac{1}{k}\sum_{l=1}^{\frac{1}{2}n_2-1}\left(\frac{\theta}{2l}\right)^k. \tag{5.3}$$

We remark that the double sum is zero if $n_2 = 2$. Similarly

$$\begin{aligned} \log\Gamma\left(\frac{n_1+\theta}{2}\right) &= \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\left\{\left(\frac{\theta}{n_1-2}\right)^k + \left(\frac{\theta}{n_1-4}\right)^k + \cdots \right. \\ &\quad \left. + \left(\frac{\theta}{2}\right)^k - s_k\left(\frac{\theta}{2}\right)^k\right\}, \end{aligned} \tag{5.4}$$

or

$$\log\Gamma\left(\frac{n_1+\theta}{2}\right) = \sum_{k=1}^{\infty}\frac{(-1)^k}{k}s_k\left(\frac{\theta}{2}\right)^k + \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\sum_{l=1}^{l=\frac{1}{2}n_1-1}\left(\frac{\theta}{2l}\right)^k. \tag{5.5}$$

By use of (5.3) and (5.5) we have for the seminvariants of z, when n_1 and n_2 are even

$$\lambda_{r:z} = \frac{(r-1)!}{2^r}\left\{\left(s_r - \sum_{k=1}^{\frac{1}{2}n_2-1}\frac{1}{k^r}\right) + (-1)^r\left(s_r - \sum_{k=1}^{\frac{1}{2}n_1-1}\frac{1}{k^r}\right)\right\}, \qquad r \geqq 2. \tag{5.6}$$

For $\lambda_{1:z} = \bar{z}$ we have by (4.6), (4.3), and (4.5)

$$\lambda_{1:z} = \frac{1}{2}\left[\left(\log n_2 - \sum_{k=1}^{\frac{1}{2}n_2-1}\frac{1}{k}\right) - \left(\log n_1 - \sum_{k=1}^{\frac{1}{2}n_1-1}\frac{1}{k}\right)\right]. \tag{5.7}$$

6. Case B, n_1 and n_2 odd. We have

$$\log \Gamma\left(\frac{n_2-\theta}{2}\right) = \log\left(\frac{n_2-2-\theta}{2}\right) + \log\left(\frac{n_2-4-\theta}{2}\right) + \cdots + \log\left(\frac{1-\theta}{2}\right) + \log\Gamma\left(\frac{1-\theta}{2}\right). \tag{6.1}$$

Expanding $\log\Gamma\left(\frac{1-\theta}{2}\right)$ by (4.5)

$$\log\Gamma\left(\frac{n_2-\theta}{2}\right) = -\left[\sum_{k=1}^{\infty}\frac{\theta^k}{k(n_2-2)^k} + \frac{\theta^k}{k(n_2-4)^k} + \cdots + \frac{\theta^k}{k}\right] + \theta\left(\sigma_1 + \frac{s_1}{2}\right) + \sum_{k=2}^{\infty}\frac{\theta^k}{k}\left(1-\frac{1}{2^k}\right)s_k. \tag{6.2}$$

However $s_k\left(1-\frac{1}{2^k}\right) = \frac{1}{1^k} + \frac{1}{3^k} + \frac{1}{5^k} + \frac{1}{7^k} + \cdots$, $k > 1$, which we shall denote hereafter by t_k. Hence (6.2) becomes

$$\log\Gamma\left(\frac{n_2-\theta}{2}\right) = \theta\left(\sigma_1+\frac{s_1}{2}\right) + \sum_{k=2}^{\infty}\frac{\theta^k}{k}t_k - \sum_{k=1}^{\infty}\frac{1}{k}\sum_{l=0}^{\frac{1}{2}(n_2-3)}\left(\frac{\theta}{2l+1}\right)^k. \tag{6.3}$$

Also

$$\log\Gamma\left(\frac{n_1+\theta}{2}\right) = \log\left(\frac{n_1+\theta-2}{2}\right) + \log\left(\frac{n_1+\theta-4}{2}\right) + \cdots + \log\left(\frac{1+\theta}{2}\right) + \log\Gamma\left(\frac{1+\theta}{2}\right), \tag{6.4}$$

and

$$\log\Gamma\left(\frac{n_1+\theta}{2}\right) = \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\left[\frac{\theta^k}{(n_1-2)^k} + \frac{\theta^k}{(n_1-4)^k} + \cdots + \frac{\theta^k}{1}\right] - \theta\left(\sigma_1+\frac{s_1}{2}\right) + \sum_{k=2}^{\infty}\frac{(-1)^k}{k}\theta^k t_k. \tag{6.5}$$

$$\log\Gamma\left(\frac{n_1+\theta}{2}\right) = -\theta\left(\sigma_1+\frac{s_1}{2}\right) + \sum_{k=2}^{\infty}\frac{(-1)^k\theta^k}{k}t_k + \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\sum_{l=0}^{l=\frac{1}{2}(n_1-3)}\frac{\theta^k}{(2l+1)^k}. \tag{6.6}$$

Combining both these results (6.3) and (6.6) we have

$$\begin{aligned}(6.7)\quad \lambda_{r:z} &= (r-1)!\left\{\left(t_r - \sum_{k=0}^{\frac{1}{2}(n_2-3)} \frac{1}{(2k+1)^r}\right)\right.\\ &\left. \qquad + (-1)^r\left(t_r - \sum_{k=0}^{\frac{1}{2}(n_1-3)} \frac{1}{(2k+1)^r}\right)\right\}, \qquad r \geqq 2.\end{aligned}$$

$$(6.8)\quad \lambda_{1:z} = \bar{z} = \left(\frac{1}{2}\log n_2 - \sum_{k=0}^{\frac{1}{2}(n_2-3)} \frac{1}{2k+1}\right) - \left(\frac{1}{2}\log n_1 - \sum_{k=0}^{\frac{1}{2}(n_1-3)} \frac{1}{2k+1}\right).$$

7. Cases C, D, and values of s_k, σ_k, t_k. The formulas for case C, n_1 even, n_2 odd are

$$(7.1)\quad \lambda_{r:z} = (r-1)!\left\{\left(t_r - \sum_{k=0}^{\frac{1}{2}(n_2-3)} \frac{1}{(2k+1)^r}\right) + \frac{(-1)^r}{2^r}\left(s_r - \sum_{k=1}^{\frac{1}{2}n_1-1} \frac{1}{k^r}\right)\right\}, \qquad r \geqq 2.$$

$$(7.2)\quad \lambda_{1:z} = \bar{z} = \frac{1}{2}\log\frac{n_2}{n_1} + \frac{1}{2}\sum_{k=1}^{\frac{1}{2}n_1-1}\frac{1}{k} - \sum_{k=0}^{\frac{1}{2}(n_2-3)} \frac{1}{2k+1} + \sigma_1.$$

The results for case D, n_1 odd, n_2 even are

$$(7.3)\quad \lambda_{r:z} = (r-1)!\left\{\frac{1}{2^r}\left(s_r - \sum_{k=1}^{\frac{1}{2}n_2-1} \frac{1}{k^r}\right) + (-1)^r\left(t_r - \sum_{k=0}^{\frac{1}{2}(n_1-3)} \frac{1}{(2k+1)^r}\right)\right\}, \qquad r \geqq 2.$$

$$(7.4)\quad \lambda_{1:z} = \bar{z} = \frac{1}{2}\log\frac{n_2}{n_1} - \sigma_1 + \sum_{k=0}^{\frac{1}{2}(n_1-3)} \frac{1}{2k+1} - \frac{1}{2}\sum_{k=1}^{\frac{1}{2}n_2-1}\frac{1}{k}.$$

We list the numerical values of s_k and t_k, $k \leqq 10$. The values of s_k are from Stieltjes [20],

$$(7.5)\quad \begin{aligned} s_1 &= 0.57721\ 56649\\ s_2 &= 1.64493\ 40668\\ s_3 &= 1.20205\ 69032\\ s_4 &= 1.08232\ 32337\\ s_5 &= 1.03692\ 77551\\ s_6 &= 1.01734\ 30620\\ s_7 &= 1.00834\ 92774\\ s_8 &= 1.00407\ 73562\\ s_9 &= 1.00200\ 83928\\ s_{10} &= 1.00099\ 45751 \end{aligned} \qquad (7.6)\quad \begin{aligned} \sigma_1 &= \log 2 = 0.69317\ 0206\\ t_2 &= 1.23370\ 00550\\ t_3 &= 1.05179\ 97903\\ t_4 &= 1.01467\ 80316\\ t_5 &= 1.00452\ 37628\\ t_6 &= 1.00144\ 70767\\ t_7 &= 1.00047\ 15487\\ t_8 &= 1.00015\ 51790\\ t_9 &= 1.00005\ 13452\\ t_{10} &= 1.00001\ 70413 \end{aligned}$$

By means of the formula $t_k = s_k\left(1 - \frac{1}{2^k}\right)$, $k > 1$, t_k was calculated from s_k. From the well known results for the Zeta function of Riemann $\zeta(s)$, [22], (p. 265, p. 267),

$$(7.7)\qquad \zeta_s = s_k = \sum_{k=1}^{\infty}\frac{1}{k^s} = \frac{1}{\Gamma(s)}\int_0^{\infty}\frac{x^{s-1}e^{-x}}{1-e^{-x}}\,dx, \qquad s \geqq 1, \qquad k > 1.$$

$$(7.8)\qquad \sigma_s = \left(1 - \frac{1}{2^{s-1}}\right)\zeta(s) = \frac{1}{\Gamma(s)}\int_0^\infty \frac{x^{s-1}}{e^x + 1}\,dx, \qquad \text{and}$$

$$(7.9)\qquad t_s = \zeta(s)\left(1 - \frac{1}{2^s}\right).$$

8. The mean of the z distribution. From our previous formulas for $\bar{z}$ we prove that if $n_1 = n_2$, $\bar{z} = 0$, and $\bar{z} < 0$ for $n_2 > n_1$, $\bar{z} > 0$ for $n_1 > n_2$. The maximum absolute value of $\lambda_{1:z}$ will occur when $n_1 = 1$, $n_2 = \infty$, or $n_1 = \infty$, $n_2 = 1$, and from (7.4) or (6.8) we have $\max|\lambda_{1:z}| = \frac{s_1}{2} + \frac{1}{2}\log 2 = .6352$.

9. Formulas for $\lambda_{2:z}$, $\mu_{2:z}$, $\lambda_{3:z}$, $\mu_{3:z}$, $\lambda_{4:z}$, and $\mu_{4:z}$. We have four cases from (5.6), (6.7), (7.1), (7.3):

$$(9.1)\qquad \begin{aligned}\lambda_{2:z} &= \frac{1}{4}\left[2s_2 - \sum_{k=1}^{\frac{1}{2}(n_1-2)}\frac{1}{k^2} - \sum_{k=1}^{\frac{1}{2}(n_2-2)}\frac{1}{k^2}\right]\\ &= .822467 - \frac{1}{4}\left(\sum_{k=1}^{\frac{1}{2}(n_1-2)}\frac{1}{k^2} + \sum_{k=1}^{\frac{1}{2}(n_2-2)}\frac{1}{k^2}\right), \qquad n_1, n_2 \text{ even.}\end{aligned}$$

$$(9.2)\qquad \lambda_{2:z} = 2.467401 - \frac{1}{4}\left(\sum_{k=0}^{\frac{1}{2}(n_2-3)}\frac{1}{(k+\frac{1}{2})^2} + \sum_{k=0}^{\frac{1}{2}(n_1-3)}\frac{1}{(k+\frac{1}{2})^2}\right), \qquad n_1, n_2 \text{ odd.}$$

$$(9.3)\qquad \lambda_{2:z} = 1.644934 - \frac{1}{4}\left(\sum_{k=1}^{\frac{1}{2}(n_1-1)}\frac{1}{k^2} + \sum_{k=0}^{\frac{1}{2}(n_2-3)}\frac{1}{(k+\frac{1}{2})^2}\right), \qquad n_1 \text{ even, } n_2 \text{ odd.}$$

$$(9.4)\qquad \lambda_{2:z} = 1.644934 - \frac{1}{4}\left(\sum_{k=0}^{\frac{1}{2}(n_1-3)}\frac{1}{(k+\frac{1}{2})^2} + \sum_{k=1}^{\frac{1}{2}(n_2-2)}\frac{1}{k^2}\right), \qquad n_1 \text{ odd, } n_2 \text{ even.}$$

In all cases of course $\lambda_{2:z} > 0$ and moreover $\lambda_{2:z} \to 0$ as n_1 and $n_2 \to \infty$. We list

$$(9.5)\qquad \lambda_{3:z} = \frac{1}{4}\left(\sum_{k=1}^{\frac{1}{2}n_1-1}\frac{1}{k^3} - \sum_{k=1}^{\frac{1}{2}n_2-1}\frac{1}{k^3}\right), \qquad n_1, n_2 \text{ even.}$$

$$(9.6)\qquad \lambda_{3:z} = \frac{1}{4}\left(\sum_{k=0}^{\frac{1}{2}(n_1-3)}\frac{1}{(k+\frac{1}{2})^3} - \sum_{k=0}^{\frac{1}{2}(n_2-3)}\frac{1}{(k+\frac{1}{2})^3}\right), \qquad n_1, n_2 \text{ odd.}$$

$$(9.7)\qquad \lambda_{3:z} = 1.803085 + \frac{1}{4}\left(\sum_{k=1}^{\frac{1}{2}(n_1-2)}\frac{1}{k^3} - \sum_{k=0}^{\frac{1}{2}(n_2-3)}\frac{1}{(k+\frac{1}{2})^3}\right), \qquad n_1 \text{ even, } n_2 \text{ odd.}$$

$$(9.8)\qquad \lambda_{3:z} = -1.803085 + \frac{1}{4}\left(\sum_{k=0}^{\frac{1}{2}(n_1-3)}\frac{1}{(k+\frac{1}{2})^3} - \sum_{k=1}^{\frac{1}{2}(n_2-2)}\frac{1}{k^3}\right), \qquad n_1 \text{ odd, } n_2 \text{ even.}$$

$$(9.9)\qquad \lambda_{4:z} = .811742 - \frac{3}{8}\left(\sum_{k=1}^{\frac{1}{2}n_2-1}\frac{1}{k^4} + \sum_{k=1}^{\frac{1}{2}n_1-1}\frac{1}{k^4}\right), \qquad n_1, n_2 \text{ even.}$$

$$(9.10)\qquad \lambda_{4:z} = 12.17614 - 6\left(\sum_{k=0}^{\frac{1}{2}(n_2-3)}\frac{1}{(2k+1)^4} + \sum_{k=0}^{\frac{1}{2}(n_1-3)}\frac{1}{(2k+1)^4}\right), \qquad n_1, n_2 \text{ odd.}$$

$$(9.11)\quad \lambda_{4:z} = 6.493939 - 6\left(\sum_{k=0}^{\frac{1}{2}(n_2-3)} \frac{1}{(2k+1)^4} + \sum_{k=1}^{\frac{1}{2}n_1-1} \frac{1}{k^4}\right), \qquad n_1 \text{ even, } n_2 \text{ odd.}$$

$$(9.12)\quad \lambda_{4:z} = 6.493939 - 6\left(\sum_{k=1}^{\frac{1}{2}n_2-2} \frac{1}{k^4} + \sum_{k=0}^{\frac{1}{2}(n_1-3)} \frac{1}{(2k+1)^4}\right), \qquad n_1 \text{ odd, } n_2 \text{ even.}$$

We see $\lambda_{r:z} > 0$ whenever r is even. If r is odd $\lambda_{r:z} < 0$ if $n_2 > n_1$, and $\lambda_{r:z} > 0$ if $n_1 > n_2$. Also $\mu_{r:z} > 0$, $n_1 > n_2$, r odd, greater than one. Similarly $\mu_{r:z} < 0$, r odd > 1, $n_2 > n_1$.

10. Skewness, excess, and values of α_n. We take for our measure of skewness $\alpha_3 = \dfrac{\mu_3}{\mu_2^{3/2}} = \dfrac{\lambda_3}{\lambda_2^{3/2}}$. For $n_2 > n_1$, $\alpha_3 < 0$. Further the skewness increases negatively if n_1 remains constant as $n_2 \to \infty$. Thus negative skewness will be a maximum for $n_2 = \infty$, $n_1 = 1$, and positive skewness will be a maximum when $n_2 = 1$, $n_1 = \infty$. The absolute value of maximum α_3 is

$$(10.1)\qquad |\alpha_3| = \left|\frac{2t_3}{t_2^{3/2}}\right| = 1.5351.$$

As our measure of kurtosis we use $\alpha_4 = \dfrac{\mu_4}{\mu_2^2} = 3 + \dfrac{\lambda_4}{\lambda_2^2}$. As a measure of excess, E, we use $E = \alpha_4 - 3 = \dfrac{\lambda_4}{\lambda_2^2}$. The excess is always positive.

11. Approximations for $\lambda_{r:z}$ by the Euler-Maclaurin sum formula. The exact results given previously for the seminvariants become unwieldy for n_1 and n_2 large. Hence we develop useful approximations for the seminvariants, and give the maximum error of the approximation. We find first our results for $\lambda_{r:z}$ when n_1 and n_2 are even and $r > 1$. We begin with (5.6)

$$\lambda_{r:z} = \frac{(r-1)!}{2^r}\left\{\left(s_r - \sum_{k=1}^{\frac{1}{2}n_2-1} \frac{1}{k^r}\right) + (-1)^r\left(s_r - \sum_{k=1}^{\frac{1}{2}n_1-1} \frac{1}{k^r}\right)\right\}$$

and rewrite this as

$$(11.1)\qquad \lambda_{r:z} = \frac{(r-1)!}{2^r}\left\{\sum_{k=\frac{1}{2}n_2}^{\infty} \frac{1}{k^r} + (-1)^r \sum_{k=\frac{1}{2}n_1}^{\infty} \frac{1}{k^r}\right\}.$$

Now find the two sums of (11.1) by the Euler-Maclaurin sum formula [21] using the first three terms, and obtain

$$(11.2)\qquad \begin{aligned} \lambda_{r:z} = \frac{(r-2)!}{2}\Bigg[&\left(\frac{n_2+r-1}{n_2^r} + (-1)^r \frac{n_1+r-1}{n_1^r}\right) \\ &+ \frac{r(r-1)}{3}\left(\frac{1}{n_2^{r+1}} + \frac{(-1)^r}{n_1^{r+1}}\right) \\ &- \frac{r(r-1)(r+1)(r+2)}{45}\left(\frac{1}{n_2^{r+3}} + \frac{(-1)^r}{n_1^{r+3}}\right)\Bigg]. \end{aligned}$$

We use the following theorem [10] (p. 539), to find the error:

If $f(x)$ is of constant sign for $x > 0$, and together with all of its derivatives, tends monotonely to zero as $x \to \infty$, Euler's summation formula may be stated in the simplified form

$$\sum_{x=0}^{n} f_x = \int_0^n f(x)\,dx + \tfrac{1}{2}(f_n + f_0) + \frac{B_2}{2!}(f_n' - f_0') + \cdots$$
$$+ \frac{(-1)^{k-1}B_{2k}}{(2k)!}(f_n^{(2k-1)} - f_0^{(2k-1)}) + \frac{\theta B_{2k+2}}{(2k+2)!}(f_n^{(2k+1)} - f_0^{(2k+1)})$$

where $0 < \theta < 1$ and $B_2 = 1/6$, $B_4 = 1/30$, $B_6 = 1/42$, $B_8 = 1/30$, $B_{10} = 5/66$, etc. If we use

$$\lambda_{r:z} = \frac{(r-2)!}{2}\left(\frac{n_2 + r - 1}{n_2^r} + (-1)^r\frac{n_1 + r - 1}{n_1^r}\right), \tag{11.3}$$

then the error committed is of the same sign and less than

$$\frac{r!}{3!}\left\{\frac{1}{n_2^{r+1}} + \frac{(-1)^r}{n_1^{r+1}}\right\}.$$

If we take

$$\lambda_{r:z} = \frac{(r-2)!}{2}\left[\left(\frac{n_2 + r - 1}{n_2^r} + (-1)^r\frac{n_1 + r - 1}{n_1^r}\right) + \frac{r(r-1)}{3}\left(\frac{1}{n_2^{r+1}} + \frac{(-1)^r}{n_1^{r+1}}\right)\right], \tag{11.4}$$

then our error is less than, and has the same sign as

$$-\frac{(r+2)!}{90}\left\{\frac{1}{n_2^{r+3}} + \frac{(-1)^r}{n_1^{r+3}}\right\}.$$

Finally if we use (11.2), our error has the same sign as, and is less than

$$\frac{(r+4)!}{945}\left\{\frac{1}{n_2^{r+5}} + \frac{(-1)^r}{n_1^{r+5}}\right\}.$$

12. Approximations for other values of n_1 and n_2, $r > 1$. Now in case n_1 and n_2 are odd we have from (6.7)

$$\lambda_{r:z} = (r-1)!\left\{\sum_{k=\frac{1}{2}(n_2-1)}^{\infty}\frac{1}{(2k+1)^r} + (-1)^r\sum_{k=\frac{1}{2}(n_1-1)}^{\infty}\frac{1}{(2k+1)^r}\right\}. \tag{12.1}$$

Applying the Euler-Maclaurin sum formula to each of the sums in (12.1) we are led to exactly the same results given in paragraph (11). The other cases are obvious combinations of the sums in (11.1) and (12.1), and so for all values of n_1 and n_2 the approximate results for $\lambda_{r:z}$, $r > 1$ are

$$\lambda_{r:z} = \frac{(r-2)!}{2}\left\{\frac{n_2+r-1}{n_2^r} + (-1)^r\frac{n_1+r-1}{n_1^r}\right\} \tag{12.2}$$
$$+\frac{r!}{6}\left\{\frac{1}{n_2^{r+1}} + \frac{(-1)^r}{n_1^{r+1}}\right\} - \frac{(r+2)!}{90}\left\{\frac{1}{n_2^{r+3}} + \frac{(-1)^r}{n_1^{r+3}}\right\}.$$

Formulas (11.1) and (12.1) prove the result previously given for $\lambda_{r:z}$ (4.7).

13. The approximate values of $\lambda_{1:z}$. From (5.7)

$$\lambda_{1:z} = \frac{1}{2}\left[\left(\log n_2 - \sum_{k=1}^{\frac{1}{2}n_2-1}\frac{1}{k}\right) - \left(\log n_1 - \sum_{k=1}^{\frac{1}{2}n_1-1}\frac{1}{k}\right)\right], \qquad n_1 \text{ and } n_2 \text{ even.}$$

We use the Euler-Maclaurin sum formula on the sum

$$\sum_{k=1}^{\frac{1}{2}n_2-1}\frac{1}{k} = \left\{\sum_{k=0}^{\frac{1}{2}n_2-1}\left(\frac{1}{k+1}\right) - \frac{2}{n_2}\right\}$$

and the similar sum involved in $\lambda_{1:z}$. Hence we have

$$\lambda_{1:z} = \frac{1}{2}\left(\frac{1}{n_2} - \frac{1}{n_1}\right) + \frac{1}{6}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right) - \frac{1}{15}\left(\frac{1}{n_2^4} - \frac{1}{n_1^4}\right), \qquad n_1, n_2 > 2. \tag{13.1}$$

The errors committed by using one, two, or three terms of (13.1) are less than, and of the same sign respectively as

$$\frac{1}{6}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right), \qquad -\frac{1}{15}\left(\frac{1}{n_2^4} - \frac{1}{n_1^4}\right), \qquad \frac{8}{63}\left(\frac{1}{n_2^6} - \frac{1}{n_1^6}\right).$$

For n_1 and n_2 both odd we find the same result as (13.1). The restriction n_1, $n_2 > 2$, may easily be replaced by $n_1, n_2 \geqq 2$ (for n_1, n_2 even) and $n_1, n_2 \geqq 1$ (for n_1, n_2 both odd). When n_1 is odd, n_2 even, the formula is again the same as (13.1) if n_1 and n_2 are sufficiently large; but if n_1 and n_2 are small we find in this case

$$\lambda_{1:z} = \frac{1}{2}\left(\frac{1}{n_2} - \frac{1}{n_1}\right) + \frac{1}{6}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right) - \frac{1}{15}\left(\frac{1}{n_2^4} - \frac{1}{n_1^4}\right) \tag{13.2}$$
$$+\frac{1}{2}\left(1 - \frac{1}{2}\right) + \frac{1}{6}\left(1 - \frac{1}{4}\right) - \frac{1}{15}\left(1 - \frac{1}{16}\right) - \frac{1}{2}\log 2.$$

Another method of finding (12.2) would have been to use the asymptotic expression for $\log \Gamma(x)$.

14. Approximate values of $\lambda_{r:z}$ for values of r. We list the approximate values of $\lambda_{r:z}$ to three terms.

$$\lambda_{1:z} = \frac{1}{2}\left(\frac{1}{n_2} - \frac{1}{n_1}\right) + \frac{1}{6}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right) - \frac{1}{15}\left(\frac{1}{n_2^4} - \frac{1}{n_1^4}\right)$$

$$
(14.1)\quad
\begin{aligned}
\lambda_{2:z} &= \frac{1}{2}\left(\frac{n_2+1}{n_2^2} + \frac{n_1+1}{n_1^2}\right) + \frac{1}{3}\left(\frac{1}{n_2^3} + \frac{1}{n_1^3}\right) - \frac{4}{15}\left(\frac{1}{n_2^5} + \frac{1}{n_1^5}\right)\\
\lambda_{3:z} &= \frac{1}{2}\left(\frac{n_2+2}{n_2^3} - \frac{n_1+2}{n_1^3}\right) + \left(\frac{1}{n_2^4} - \frac{1}{n_1^4}\right) - \frac{4}{3}\left(\frac{1}{n_2^6} - \frac{1}{n_1^6}\right)\\
\lambda_{4:z} &= \left(\frac{n_2+3}{n_2^4} + \frac{n_1+3}{n_1^4}\right) + 4\left(\frac{1}{n_2^5} + \frac{1}{n_1^5}\right) - 8\left(\frac{1}{n_2^7} + \frac{1}{n_1^7}\right)\\
\lambda_{5:z} &= 3\left(\frac{n_2+4}{n_2^5} - \frac{n_1+4}{n_1^5}\right) + 20\left(\frac{1}{n_2^6} - \frac{1}{n_1^6}\right) - 56\left(\frac{1}{n_2^8} - \frac{1}{n_1^8}\right)\\
\lambda_{6:z} &= 12\left(\frac{n_2+5}{n_2^6} + \frac{n_1+5}{n_1^6}\right) + 120\left(\frac{1}{n_2^7} + \frac{1}{n_1^7}\right) - 448\left(\frac{1}{n_2^9} + \frac{1}{n_1^9}\right).
\end{aligned}
$$

The approximate values given by Cornish and Fisher [8] (p. 319), are similar, but have fewer terms. Cornish and Fisher give no remainder term. From (14.1) and (12.2) we see the maximum absolute values of $\lambda_{2r+1:z}$, $r \geqq 1$, occur when $n_2 = \infty$, $n_1 = 1$, or $n_2 = 1$, $n_1 = \infty$. Similarly $\lambda_{2r:z}$, $r \geqq 1$, has its maximum value for $n_1 = n_2 = 1$. The standard seminvariants of z are defined $\xi_{r:z} = \dfrac{\lambda_r}{\lambda_2^{\frac{1}{2}r}}$, $r \geqq 2$. We also note that for $n_2 > n_1$, $\xi_{2r+1:z} < 0$, $r \geqq 1$ and hence $\alpha_{2r+1} < 0$ also where $\alpha_n = \dfrac{\mu_n}{\mu_2^{\frac{1}{2}n}}$. Moreover the maximum absolute values of $\xi_{2r:z}$ and $\xi_{2r+1:z}$ occur when $n_1 = 1$, $n_2 = \infty$ or $n_2 = 1$, $n_1 = \infty$; and also for α_{2r} and α_{2r+1}. Approximately then

$$
(14.2)\qquad \max \xi_{r:z} = (-1)^r \frac{(r-1)!}{2}, \qquad r \geqq 2.
$$

The exact value for maximum $\alpha_{4:z}$ is $3 + \dfrac{6t_4}{t_2^2} = 7.07$.

15. Approach to normality of the z distribution. We prove the theorem: The distribution of z approaches normality as n_1 and $n_2 \rightarrow \infty$ in any manner whatever, with $\bar{z} = \dfrac{1}{2}\left(\dfrac{1}{n_2} - \dfrac{1}{n_1}\right)$, $\sigma_z^2 = \dfrac{1}{2}\left(\dfrac{1}{n_2} + \dfrac{1}{n_1}\right)$. We also find an upper bound of the absolute value of the difference between the z distribution and the function determined by the approximate seminvariants of z when n_1 and n_2 become large. To prove the theorem we start with the original distribution of z, and find when n_1 and n_2 are large,

$$
(15.1)\qquad P(z) = \frac{1}{\sqrt{2\pi}\,\sigma_z}\left\{\frac{n_1+n_2}{n_1e^{2z}+n_2}\right\}^{\frac{1}{2}(n_1+n_2)} e^{n_1 z}\,dz.
$$

We change to standard units $z = t\sigma_z + \bar{z}$, then

$$P(t) = \frac{1}{\sqrt{2\pi}} \left\{ \frac{n_1 + n_2}{n_1 e^{2t\sigma + 2\bar{z}} + n_2} \right\}^{\frac{1}{2}(n_1+n_2)} e^{n_1 t\sigma + n_1 \bar{z}}\, dt, \qquad -\infty < t < \infty. \tag{15.2}$$

We rewrite this as

$$P(t) = \frac{1}{\sqrt{2\pi}} \left\{ \frac{n_1 + n_2}{n_1 e^{2n_2(t\sigma+\bar{z})/(n_1+n_2)} + n_2 e^{-2n_1(t\sigma+\bar{z})/(n_1+n_2)}} \right\}^{\frac{1}{2}(n_1+n_2)} dt. \tag{15.3}$$

Expand $n_1 e^{2n_2(t\sigma+\bar{z})/(n_1+n_2)}$ and $n_2 e^{-2n_1(t\sigma+\bar{z})/(n_1+n_2)}$ and add term by term. Divide this result by $n_1 + n_2$ from the numerator of $P(t)$ to obtain

$$1 + \frac{2n_1 n_2 (t\sigma + \bar{z})^2}{(n_1 + n_2)^2} + O_1 \left\{ \frac{1}{(n_1 + n_2)^{\frac{3}{2}}} \right\}. \tag{15.4}$$

Hence

$$P(t) = \frac{1}{\sqrt{2\pi}} \left\{ 1 + \frac{2n_1 n_2 (t\sigma + \bar{z})^2}{(n_1 + n_2)^2} \right\}^{-\frac{1}{2}(n_1+n_2)} dt. \tag{15.5}$$

We evaluate (15.5) for n_1 and n_2 large by using logarithms.

$$-\frac{n_1 + n_2}{2} \log \left\{ 1 + \frac{2n_1 n_2 (t\sigma + \bar{z})^2}{(n_1 + n_2)^2} \right\}$$
$$= -\frac{n_1 + n_2}{2} \left[\left\{ \frac{2n_1 n_2 (t\sigma + \bar{z})^2}{(n_1 + n_2)^2} \right\} - \frac{1}{2} \left\{ \frac{2n_1 n_2 (t\sigma + \bar{z})^2}{(n_1 + n_2)^2} \right\}^2 + \sum_{r=3}^{\infty} \frac{(-1)^{r+1}}{r} \left\{ \frac{2n_1 n_2 (t\sigma + \bar{z})^2}{(n_1 + n_2)^2} \right\}^r \right].$$

This gives

$$-\frac{\sigma^{-2}}{2} (t^2\sigma^2 + 2t\sigma\bar{z} + \bar{z}^2) + \frac{n_1^2 n_2^2}{(n_1 + n_2)^3} (t\sigma + \bar{z})^4 + \sum_{r=3}^{\infty} (-1)^r \frac{\{2n_1 n_2 (t\sigma + \bar{z})^2\}^r}{2r(n_1 + n_2)^{2r-1}}.$$

We reduce this then to

$$-\frac{t^2}{2} - \sigma^{-1}\bar{z}t - \frac{(\bar{z}\sigma^{-1})^2}{2} + \frac{1}{2} \left\{ \frac{2n_1^2 n_2^2}{(n_1 + n_2)^2} \right\} \frac{(t\sigma + \bar{z})^4}{n_1 + n_2}$$

+ terms involved in the above summation. Let $U = \sigma^{-1}\bar{z} < \sigma$. Since $\lim_{n_1, n_2 \to \infty} \sigma = 0$, $\lim_{n_1, n_2 \to \infty} U = 0$. Similarly $\lim_{n_1, n_2 \to \infty} \frac{\bar{z}^2 \sigma^{-2}}{2} = \lim_{n_1, n_2 \to \infty} \frac{U^2}{2} = 0$. Consider $\frac{n_1^2 n_2^2}{(n_1 + n_2)^3} (t\sigma + \bar{z})^4 = \frac{\sigma^{-4}(t\sigma + \bar{z})^4}{4(n_1 + n_2)} = \frac{(t + U)^4}{4(n_1 + n_2)}$. Hence $\lim_{n_1, n_2 \to \infty} \frac{(t + U)^4}{4(n_1 + n_2)} = 0$. In like fashion

$$\sum_{r=3}^{\infty} \frac{(-1)^r}{2r} \left\{ \frac{2n_1 n_2}{n_1 + n_2} \right\}^r \frac{(t\sigma + \bar{z})^{2r}}{(n_1 + n_2)^{r-1}} = \sum_{r=3}^{\infty} \frac{(-1)^r \sigma^{-2r} (t\sigma + \bar{z})^{2r}}{2r(n_1 + n_2)^{r-1}}.$$

Now clearly from our previous discussion for $r = 2$, we see

$$\lim_{n_1,n_2\to\infty} \sum_{r=3}^{\infty} \frac{(-1)^r}{2r} \frac{\sigma^{-2r}(t\sigma + \bar{z})^{2r}}{(n_1+n_2)^{r-1}} = 0.$$

This completes the proof.

We now consider the function, $f(z)$, determined by the approximate seminvariants of z. We start with

$$\lambda_{1:z} = \frac{1}{2}\left(\frac{1}{n_2} - \frac{1}{n_1}\right) \quad \text{and} \quad \lambda_{r:z} = \frac{(r-2)!}{2}\left\{\frac{n_2+r-1}{n_2^r} + (-1)^r \frac{n_1+r-1}{n_1^r}\right\}, \quad r > 1,$$

from (12.2) using only the first term. We may easily prove then that as n_1 and n_2 approach infinity in any manner whatever the function $f(z)$ represents a normal frequency distribution with

$$\bar{z} = \frac{1}{2}\left(\frac{1}{n_2} - \frac{1}{n_1}\right) \quad \text{and} \quad \mu_{2:z} = \frac{1}{2}\left(\frac{n_2+1}{n_2^2} + \frac{n_1+1}{n_1^2}\right).$$

This further shows the identity of $f(z)$ and $y(z)$ in the limit as n_1 and $n_2 \to \infty$.

Since the moment generating function of $f(z)$ is

$$\left(1 - \frac{\theta}{n_2}\right)^{\frac{1}{2}(n_2-1-\theta)} \left(1 + \frac{\theta}{n_1}\right)^{\frac{1}{2}(n_1-1+\theta)}$$

we have

$$f(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\theta z}\left(1 - \frac{i\theta}{n_2}\right)^{\frac{1}{2}(n_2-1-i\theta)} \left(1 + \frac{i\theta}{n_1}\right)^{\frac{1}{2}(n_1-1+i\theta)} d\theta. \tag{15.6}$$

I have not been able to evaluate (15.6). We instead shall find an upper bound to the difference $| f(z) - y(z) |$ as n_1 and n_2 become large. We form $f(z) - y(z)$. Then by use of Stirling's formula for $n!$ with the remainder term and by the Fourier Integral Theorem,

$$| f(z) - y(z) | \leqq (e^{\beta_3/6n_1 + \beta_4/6n_2} - 1)y(z) \text{ where } 0 < \beta_3 < 1, 0 < \beta_4 < 1, \tag{15.7}$$

and

$$\lim_{n_1,n_2\to\infty} | f(z) - y(z) | = 0, \text{ and for this case } f(z) = y(z). \tag{15.8}$$

Of course (15.7) furnishes the upper bound of the absolute value between the frequency distribution of z and the function determined by the approximate seminvariants of z for any values of n_1 and n_2.

Up to this point we have assumed that there exists a function determined by the seminvariants

$$\lambda_{1:z} = \frac{1}{2}\left(\frac{1}{n_2} - \frac{1}{n_1}\right) \quad \text{and} \quad \lambda_{r:z} = \frac{(r-2)!}{2}\left\{\frac{n_2+r+1}{n_2^r} + (-1)^r \frac{n_1+r-1}{n_1^r}\right\}, \quad r > 1.$$

This may readily be proved by using the following theorem [18] (p. 536): The

determined character of the moments problem for an infinite interval is insured if $\sum_{n=1}^{\infty} c_{2n}^{-1/2n}$ diverges $\left(c_n = \int_{-\infty}^{\infty} x^n \, dF(x)\right)$.

16. The Pearson types of approximating curve. In discussing the types of the Pearson system which may be expected to approximate the z distribution we shall use the results of H. C. Carver [1], and the further exposition of C. C. Craig [3]. To find the Pearson type we compute $\delta = \dfrac{2\alpha_4 - 3\alpha_3^2 - 6}{\alpha_4 + 3}$. We shall find it convenient to use the approximations $\alpha_3 = \dfrac{\sqrt{2}\,(n_1 - n_2)}{\sqrt{n_1 n_2 (n_1 + n_2)}}$ and $\alpha_4 = 3 + 4\,\dfrac{(n_1^2 - n_1 n_2 + n_2^2)}{n_1 n_2 (n_1 + n_2)}$ to obtain

$$\delta = \frac{(n_1 + n_2)^2}{3n_1^2 n_2 + 3n_1 n_2^2 + 2n_1^2 - 2n_1 n_2 + 2n_2^2}, \tag{16.1}$$

and consequently $0 < \delta \leqq \frac{1}{2}$. The only possibilities are Types IV, VII, VI, or V since the greatest value of α_3^2 by (14.1) is 2.3565. Now if $n_1 = n_2$, we have Type VII, since $\alpha_3 = 0$, $\delta > 0$. In all other cases we shall have Types IV, V, or VI according as $\alpha_3^2 < 4\delta(\delta + 2)$, $\alpha_3^2 = 4\delta(\delta + 2)$, $\alpha_3^2 > 4\delta(\delta + 2)$. We neglect δ^2. Hence $\alpha_3^2 < 8\delta$ implies

$$\begin{aligned} n_2^4(n_1 - 2) + n_2^3(15n_1^2 + 6n_1) + n_2^2(15n_1^3 - 8n_1^2)& \\ + n_2(n_1^4 + 6n_1^3) - 2n_1^4 > 0.& \end{aligned} \tag{16.2}$$

A simple investigation reveals then the following results:

$$\begin{aligned} &\text{Type IV for } n_1, n_2 \geqq 2,\ n_1 \neq n_2. \\ &\text{Type IV for } n_1 = 1,\ 1 \leqq n_2 \leqq 21; \text{ or } n_2 = 1,\ 1 \leqq n_1 \leqq 21. \\ &\text{Type VI for } n_1 = 1,\ n_2 > 22. \\ &\qquad\quad\ \text{for } n_2 = 1,\ n_1 > 22. \\ &\text{Type VII for } n_1 = n_2. \end{aligned} \tag{16.3}$$

Clearly the z distribution has features comparable to Type IV since both have infinite range. However, Type IV is irksome to fit in practice.

17. The Type III approximating curve, the logarithmic curve, and the Gram-Charlier Type A. The criterion for Type III is $\delta = 0$, $\alpha_3 \neq 0$. We see that as n_1 and n_2 increase the value of δ will decrease. Even for small values of n_1 and n_2 Type III will furnish a fair approximation to the z distribution. For example $n_1 = 10$, $n_2 = 5$, $\delta = .094$. The advantage of the Type III approxi-

mation rests on the fact that Salvosa's tables may be used. From the chart in [16] since $\alpha_3^2 \leqq 2.3565$, we are assured that the approximating Type III curve is bell shaped. For $n_1 = 1, 2$, $n_2 =$ any value, this approximation is not all that could be desired, although even in such cases it does have value. We note that Type III has limited range at one extreme $\left(-\frac{2}{\alpha_3}, \infty\right)$ while the range of the z distribution is $(-\infty, \infty)$. Salvosa's tables extend as far as $\alpha_3 = 1.1$, and since max $\alpha_3 = 1.5351$, we see in some cases, and these only for $n_1 = 1$, n_2 large, we shall be obliged to make use of Pearson's *Tables of the Incomplete Gamma Function* [14]. The logarithmic frequency curve

$$f(u) = \frac{1}{\sqrt{2\pi}\, c(u-a)} \exp\left[-\frac{1}{2c^2}\left(\log\frac{u-a}{b}\right)^2\right]$$

will be useful in approximating the z distribution. While it has been discussed by many authors we shall follow Pae-Tsi Yuan [23], where a full bibliography may be found. In our discussion we use the $\beta_1 = \alpha_3^2$, $\beta_2 = \alpha_4$ chart of the Pearson system as given by S. J. Pretorius [16] (p. 147), since the logarithmic frequency locus connecting α_3^2 and α_4 is already drawn in. The justification of this curve for fitting is due to the fact that in the β_1, β_2 chart of the Pearson system as given by S. J. Pretorius [16] (p. 147), the logarithmic frequency locus lies in the Type VI region between the Type III locus and the Type V locus, and consequently closer to the Type IV region than Type III itself does. Hence since Type III fits fairly well under certain conditions and Type IV fits well we can expect the same for the logarithmic curve. Furthermore when α_3 is small the logarithmic curve is similar to Type III [23] (p. 42), and as α_3 becomes larger, $\alpha_3 = 1$, the difference between the two types is pronounced. However, it is just when α_3 becomes large in the region $n_1 = 1$, $n_2 \geqq 22$ that we find the logarithmic curves give a fine fit, since in such cases the point (α_3^2, β_2) lies practically on the logarithmic locus [16]. To fit the curve [23] (pp. 37, 48, 49), we find the values of the three parameters a, b, c. To find c we solve the equation $w^3 + 3w^2 - (4 + \alpha_{3:z}^2) = 0$ for w using the table [23] (p. 48) given by Pae-Tsi Yuan. Knowing w we can easily solve for

$$\text{(17.1)} \qquad \begin{aligned} c &= (\log w)^{\frac{1}{2}}, & b &= \left(\frac{w+2}{\alpha_{3:z}}\right) w^{-\frac{1}{2}} \sigma_z, \\ a &= \bar{z} - \frac{(w+2)\sigma_z}{\alpha_{3:z}}, & t &= \frac{z-\bar{z}}{\sigma_z} = \frac{e^{xc-\frac{1}{2}c^2} - 1}{(e^{c^2}-1)^{\frac{1}{2}}} \end{aligned}$$

where the value of x must be obtained from the table of areas under the normal curve, if the z distribution is approximated by use of areas.

Since the Gram-Charlier Type A series generally approximates a Pearson Type IV fairly well when α_3^2 is not too large, it is to be expected that the Type A series will approximate the z distribution in those cases when $n_1 = n_2$, and also when α_3^2 is not too large.

18. Levels of significance and approximation methods. We shall apply the results of the previous paragraphs to the determination of the value of z for any level of significance α, i.e. the value of z such that $\int_{-\infty}^{z} y(z)\, dz = 1 - \alpha$. We have such levels as the median (the 50% point of significance), the 20%, 5%, 1%, and .1% points as given in [9]. Where these tables apply there is no need for other methods. It would be desirable to extend the results for any level of significance whatever. The methods which we shall use are (1) the logarithmic frequency curve, (2) the Gram-Charlier Type A, and (3) the Type III approximation. For finding the levels of significance by the Incomplete Beta function, the reader is referred to [13], (p. lviii, topic (viii)). The logarithmic curve is very simple to use in conjunction with the table of areas under the normal curve. From Pae-Tsi Yuan we have

$$t = \frac{e^{xc - \frac{1}{2}c^2} - 1}{(e^{c^2} - 1)^{\frac{1}{2}}}, \quad \text{where } (e^{c^2} - 1)^{\frac{1}{2}} \tag{18.1}$$

takes the same sign as α_3. The value of x is obtained from the table of the normal curve, 1.64 for the 5% level, 2.33 for the 1% level; the value of c is obtained from w (17.1), and consequently the value of t (18.1). Then we have if z_α = value of z for any level of significance, $t = \dfrac{z_\alpha - \bar{z}}{\sigma_z}$ to solve for z_α, where $\bar{z}$, and σ_z are the values of the mean and standard deviation of z as given by the proper formulas in (5), (6), (7). We illustrate with examples:

(18.2) 5% point of z, $n_1 = \infty$, $n_2 = 1$. $\alpha_3 = 1.5351$, $w = 1.2264$, $x = 1.64$, $t = 1.88$, $\bar{z} = .6352$, $\sigma_z = 1.11$, and as a result $z_{5\%} = 2.72$. Fisher [9] gives 2.7693.

We can also find $z_{5\%}$ easily for $n_1 = 1$, $n_2 = \infty$. Here $\alpha_3 = -1.5351$, $w = 1.2264$, $x = -1.64$, $t = 1.197$, $\bar{z} = -.6352$, $\sigma_z = 1.11$, $z_{5\%} = .694$ compared with Fisher [9] $z_{5\%} = .6729$.

(18.3) 1% point for $n_1 = 4$, $n_2 = 8$, $\bar{z} = -.0701$, $\sigma_z = .4819$, $\alpha_{3:z} = -.3619$, $w = 1.0144$, $t = 2.17$ and $z_{1\%} = .976$, while the accurate result is .9734.

From experience the values of z for any level of significance obtained by the logarithmic frequency curve will possess an error less than 2% of the true value of z for the level of significance if n_1 and n_2 are greater than twenty. It would seem that for other values of n_1 and n_2 the error could not be greater than 10%, and usually would be much less.

19. The Gram-Charlier Type A. We take the series in the form

$$F(t) = \varphi(t) + A_3\varphi^{(3)}(t) + A_4\varphi^{\text{IV}}(t), \qquad \varphi(t) = \frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$$

$$t = \frac{z - \bar{z}}{\sigma_z}, \qquad A_3 = \frac{-\lambda_{3:t}}{3!}, \qquad A_4 = \frac{\lambda_{4:t}}{4!}.$$

Some examples follow.

(19.1) We use the material of (18.3) and employ three terms of $F(t)$. $\bar{z} = -.0701$, $\sigma_z = .4819$, $\lambda_{3:z} = -.0405$, $\lambda_{4:z} = .0336$, $A_3 = .06032$, $A_4 = .02596$.

Fitting $F(t)$ by ordinates we have $t = 2.17$, and consequently $z = .976$.

(19.2) We take $n_1 = n_2 = 5$, $\bar{z} = 0$, $\sigma_z = .4952$, $\lambda_{3:z} = 0$, $\lambda_{4:z} = .02798$, $A_3 = 0$, $A_4 = .01939$.

5% point: By ordinates $t = 1.57$, $z_{5\%} = .777$, while Fisher gives .8097.
1% point: By ordinates $t = 2.325$, $z_{1\%} = 1.15$, while Fisher gives 1.1974.

(19.3) We take $n_1 = 3$, $n_2 = 20$, $\bar{z} = -.15909$, $\sigma_z = .5099$, $\lambda_{3:z} = -.10222$, $\lambda_{4:z} = .08822$, $A_3 = .12854$, $A_4 = .05438$. By ordinates $t = 1.523$, $z_{5\%} = .618$, Fisher gives .5654. $t = 1.989$, $z_{1\%} = .855$, Fisher gives .7985. The Gram-Charlier Type A is recommended only for $n_1 = n_2$ and $n_1, n_2 \geqq 20$.

20. Type III approximation, the median, and 5% point. Since for Type III the median, m_z, is approximately two-thirds of the distance from the mode to the median if α_3 is moderate [12], [6], then we have further assuming n_1, $n_2 \geqq 20$.

$$m_z = \frac{1}{3}\left(\frac{1}{n_1} - \frac{1}{n_2}\right) + \frac{1}{9}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right). \tag{20.1}$$

From experience this result will furnish an accuracy with an error less than 2% of the true value in the range above indicated.

$$t_{5\%} = 1.6437 + .2760\alpha_3 - .04506\alpha_3^2. \tag{20.2}$$

This was found by use of Salvosa's tables and for $\alpha_3 > 1.1$ by [14].

$$z_{5\%} = \sigma_z[1.644 + .2760\alpha_{3:z} - .0451\alpha_{3:z}^2] + \bar{z}. \tag{20.3}$$

We illustrate the use of (20.3) with some examples.

(20.4) $n_1 = n_2 = 1$, $\sigma_z = 1.5706$, $\alpha_{3:z} = 0$, $\bar{z} = 0$, $z_{5\%} = 2.582$,

while the accurate value is $z_{5\%} = 2.5421$.

(20.5) $n_1 = \infty$, $n_2 = 1$, $\alpha_3 = 1.5351$, $\bar{z} = .6352$, $\sigma_z = 1.11$, $z_{5\%} = 2.81$. The accurate value is 2.7693.

(20.6) $n_1 = n_2 = 5$, $\sigma_z = .4952$, $\alpha_{3:z} = 0$, $\bar{z} = 0$, $z_{5\%} = .8141$, while the accurate value is $z_{5\%} = .8097$.

(20.7) $n_1 = 4$, $n_2 = 8$, $\bar{z} = -.0701$, $\sigma_z = .4819$, $\alpha_3 = -.3619$, $z_{5\%} = .6712$, while the accurate value is .6725.

(20.8) $n_1 = 1$, $n_2 = 10$, $\bar{z} = -.5835$, $\sigma_z = 1.1353$, $\alpha_3 = -1.4333$, $z_{5\%} = .7283$, while the accurate value is .8012.

In a future paper exactly the same methods will be used for any per cent point of z whatever in order to compare with the results of W. G. Cochran [2]. If

n_1 and n_2 are large we may use the approximate formulas for σ_z, $\alpha_{3:z}$, and $\bar{z}$ to obtain to the order of σ_z^3,

$$(20.9) \qquad z_{5\%} = 1.644\sigma_z + .7760\left(\frac{1}{n_2} - \frac{1}{n_1}\right), \quad \text{where} \quad \sigma_z = \sqrt{\frac{1}{2}\left(\frac{1}{n_2} - \frac{1}{n_1}\right)}.$$

We expand Fisher's result [9]
$z_{5\%} = \dfrac{1.6449}{\sqrt{n-1}} + .7843\left(\dfrac{1}{n_2} - \dfrac{1}{n_1}\right)$ by the binomial theorem, where $h = \dfrac{1}{\sigma_z^2}$, to obtain a comparable result

$$(20.10) \qquad z_{5\%} = 1.645\sigma_z + .7843\left(\frac{1}{n_2} - \frac{1}{n_1}\right).$$

The numerical examples given in this chapter illustrate unfavorable cases as well as favorable ones.

21. The distribution of F. Historically Snedecor [19] was the first to use F for e^{2z}. We find

$$(21.1) \qquad P(F) = \frac{n_1^{\frac{1}{2}n_1} n_2^{\frac{1}{2}n_2}}{\mathrm{B}\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \frac{F^{\frac{1}{2}n_1 - 1}}{(n_1 F + n_2)^{\frac{1}{2}(n_1+n_2)}} dF, \qquad 0 \leqq F \leqq \infty.$$

The distribution of F is J shaped if $n_1 \leqq 2$, and bell shaped for $n_1 > 2$, and for $n_1 > 2$ one mode exists, $F_0 = \dfrac{n_2(n_1 - 2)}{n_1(n_2 + 2)}$. The two points of inflection, which exist for $n_1 \geqq 4$, are equidistant from the mode. The moments are

$$\mu_m' = \left(\frac{n_2}{n_1}\right)^m \frac{\Gamma\left(\frac{n_1 + 2m}{2}\right)\Gamma\left(\frac{n_2 - 2m}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}, \qquad n_2 > 2m$$

$$\bar{F} = \frac{n_2}{n_2 - 2}, \qquad n_2 > 2, \qquad \mu_2 = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)} \sim 2\left(\frac{1}{n_1} + \frac{1}{n_2}\right),$$

$$\alpha_{3:F} \sim \frac{2\sqrt{2}(2n_1 + n_2)}{\sqrt{n_1 n_2(n_1 + n_2)}}.$$

The exact results for μ_3, μ_4, α_3, and α_4 are omitted because of length. We have the theorem that as n_1, $n_2 \to \infty$ in any manner whatever the distribution of F approaches normality with mean $\bar{F} = 1$, $\sigma_F = \sqrt{2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$. The proof is omitted. The only type of approximating curve of any value is Type III. Of course the distribution of F is Type VI. No tables exist for Type VI. Furthermore the F distribution approaches the Type III function so slowly as to make most approximations of little value unless $\alpha_{3:F} \leqq 1.1$. Other possible

parameters are $\theta = \frac{n_1(n_2 + 1)}{n_2(n_1 + 1)} F$, and $H = \frac{n_1 F}{n_2 + n_1 F}$, [13]. Since $| \alpha_{3:H} | = 2 | \alpha_{3:z} |$ approximately we see that the distribution of H is more skewed than that of z. We mention briefly also $S_1^2 - S_2^2$ where $S_1^2 = \frac{n_1}{N_1} s_1^2$, $S_2^2 = \frac{n_2}{N_2} s_2^2$. Clearly z, F, θ, and H give equivalent levels of significance. This is not true for z and $S_1^2 - S_2^2$.

Finally, since $F = \frac{s_1^2}{s_2^2}$, it may be interpreted as a quotient [5]. When the moments of F do not exist, it is due to the distribution function of s_2^2.

22. Conclusion. We have found the seminvariants for the z distribution, and approximations for them. Type III, and the logarithmic normal frequency functions are shown to be excellent approximations to the z distribution. The approach to normality for the z distribution is proved. A formula is given for finding the 5% level of significance for z. The F distribution is studied along the same lines. As far as the construction of tables for levels of significance is concerned, the z distribution is much easier to use. My sincerest thanks are due Professor C. C. Craig for his helpful guidance and many suggestions.

BIBLIOGRAPHY

[1] H. C. Carver, *Handbook of Mathematical Statistics*, H. L. Rietz, ed., Boston: Houghton-Mifflin Co., 1924. Chapter on frequency curves.

[2] W. G. Cochran, "Note on an approximate formula for the significance levels of z," *Annals of Math. Stat.*, Vol. 11 (1940), pp. 93–95.

[3] C. C. Craig, "A new exposition and chart for the Pearson system of frequency curves," *Annals of Math. Stat.*, Vol. 7 (1936), pp. 16–28.

[4] C. C. Craig, "An application of Thiele's semi-invariants to the sampling problem," *Metron*, Vol. 7 (1928–29), pp. 3–74.

[5] C. C. Craig, "The frequency function of y/x," *Annals of Math.*, Second Series, Vol. 30 (1929), pp. 471–486.

[6] A. T. Doodson, "Relation of a mode, median, and mean in a frequency curve," *Biometrika*, Vol. 11 (1917), p. 425.

[7] R. A. Fisher, "On a distribution yielding the error functions of several well known statistics," *Proc. International Math. Cong.*, 1924, Toronto, Vol. 2, pp. 805–813.

[8] R. A. Fisher, and E. A. Cornish, "Moments and cumulants in the specification of distributions," *Revue de l'Institut International de Statistics*, 5th year, pp. 307–20, 1937, La Hague.

[9] R. A. Fisher, and Yates, *Statistical Tables for Biological, Agricultural, and Medical Research*, London: Oliver and Boyd, 1938.

[10] K. Knopp, *Theory and Application of Infinite Series*, English translation, Edinburgh: Blackie and Son, 1928.

[11] N. Nielsen, *Handbuch der Theorie der Gamma Functionen*, Leipzig: Teubner, 1906.

[12] C. A. Olshen, "Transformation of the Pearson Type III distribution," *Annals of Math. Stat.*, Vol. 9 (1938), pp. 176–200.

[13] K. Pearson (Editor), *Tables of the Incomplete Beta Function*, London: Biometrika Office, University College, London, 1934.

[14] K. Pearson (Editor), *Tables of the Incomplete Gamma Function*, London: His Majesty's Stationery Office, 1922.

[15] K. PEARSON, S. A. STOUFFER, and F. N. DAVID, "Further applications in statistics of the $T_m(x)$ Bessel function," *Biometrika*, Vol. 24 (1932), pp. 293–350.

[16] S. J. PRETORIUS, "Skew bivariate frequency curves examined in the light of numerical illustrations," *Biometrika*, Vol. 22, (1930–31).

[17] L. R. SALVOSA, "Tables of Pearson's Type III function," *Annals of Math. Stat.*, Vol. 1 (1930), pp. 191–8 et seq.

[18] J. SHOHAT, and M. FRECHET, "A proof of the generalized second limit-theorem in the theory of probability," *Trans. Am. Math. Soc.*, Vol. 33 (1931), pp. 531–43.

[19] G. W. SNEDECOR, *Calculation and Interpretation of the Analysis of Variance and Covariance*, Ames, Iowa: Collegiate Press.

[20] T. J. STIELTJES, "Tables des valeurs des sommes $s_k = \sum_{n=1}^{\infty} n^{-k}$," *Acta Math.*, Vol. 10, pp. 299–302.

[21] WHITTAKER and ROBINSON, *The Calculus of Observations*, Edinburgh: Blackie and Son, second edition, p. 135.

[22] WHITTAKER and WATSON, *Modern Analysis*, 4th edition, London: Cambridge University Press, 1935.

[23] PAE-TSI YUAN, "On the logarithmic frequency distribution and the semi-logarithmic correlation surface," *Annals of Math. Stat.*, Vol. 4, (1933).

[24] R. T. ZOCH, "Some interesting features of frequency curves," *Annals of Math. Stat.*, Vol. 4, (1935), pp. 1–10.

THE DOOLITTLE TECHNIQUE

BY PAUL S. DWYER

University of Michigan

1. Introduction. Most authors who have presented the Doolittle method, from Doolittle [1] down to the present, have not given any formal proof that the solution is valid in the general case. They usually are content with a form describing the various steps of a Doolittle solution.

The author has recently shown [2] that the Doolittle method can be abbreviated to a technique which is also an abbreviation, essentially, of the method of single division and its abbreviation which Aitken called the "Method of Pivotal Condensation" [3]. It appears at once that the validity of the Doolittle method follows from the validity of the method of single division—a validity which is readily established.

However one may desire a "proof" which is based directly on the Doolittle technique without referring to other methods of solution. It is the chief purpose of this paper to present such a proof. It is accomplished by the introduction of a notation which precisely describes the conventional Doolittle process and by proving that this process results in a system of equations whose prediagonal terms are zero. It is a secondary purpose of the paper to emphasize the advantages of the Abbreviated Doolittle method and to explain and illustrate minor variations in the conventional Doolittle technique.

2. The Abbreviated Doolittle solution. We first direct our attention to the essential parts of a Doolittle solution and these are the last two rows of each matrix of the standard Doolittle presentation. The additional rows in the standard presentation are rows of products which are used solely for the purpose of finding the two bottom rows of each matrix and they need not be recorded, if a computing machine is available, since the essential information is present in the two bottom rows. Doolittle [1] did not have calculating machines (he used multiplication tables) but he put the important information in Table A and carefully segregated the supplementary information in Table B. With reference to this he wrote [1]

"It is to be observed that the numbers in Table B have but a single use while those in Table A are used over and over, and where the number of equations is large, it is of great advantage that they should be thus tabulated by themselves in a form compact and easy of reference."

For purposes of proof, as well as for purposes of calculation if a computing machine is available, it is only necessary to utilize the forward part of the Abbreviated Doolittle solution which is the equivalent of the Doolittle Table A.

A four variable illustration of the Abbreviated Doolittle technique is presented in Table I. The successive equations are indicated by number, as is customary, and the operation which defines the equation is specified. The actual operation is indicated more explicitly by the notation of column 3 and this is discussed in the next section.

The presentation of Table I introduces one variation from the standard Doolittle method. The division is made by the diagonal coefficient of each row rather than by its negative. One may still use the old technique, if he prefers, but it is felt that one can subtract products as easily as he can add products with modern machines equipped with automatic negative multiplication. In addition the entries of the equivalent rows then have the same signs and, too, it is not necessary to take the time to change the signs of the second rows. This variation uses the same division method as the method of single division [2] and as the method of pivotal condensation [3] so that the abbreviated form of these methods is, essentially, the same as the abbreviated form of the Doolittle method.

The application of this technique leads at each step to a coefficient for each variable. However if the process is to lead from our four equations in four unknowns, to three in three, to two in two, to one in one, it follows that all the entries to the left of the diagonal, which we may call prediagonal entries, must be zero. That this is true in the general case is the objective of the proofs of later sections.

3. A notation for and description of the Doolittle technique. A main contribution of the present article is the use of a notation which describes the Doolittle technique. As long as the Doolittle process is described loosely by means of "operations" it is difficult to be precise in defining quantities which appear in the calculation, but when a notation is used which is definite enough to permit expansion in terms of the original coefficients, some sort of proof may be available. The present notation bears some resemblance to that suggested by Gauss [4], though Gauss used letters to indicate the primary subscripts and numbers to indicate the number of secondary subscripts and his notation was directly applicable to the sums of least squares theory rather than to symmetric equations in general.

We wish to find the solution of the equations

$$\sum_{i=1}^{n} a_{ij}x_i = a_{n+1,j}, \qquad j = 1, 2, \cdots, n \tag{1}$$

where the matrix of the coefficients is symmetric. We do this by obtaining auxiliary equations which feature a decreasing number of variables. No serious restriction is made if we assume that the variables x_1, x_2, x_3, etc., are eliminated successively. The Doolittle technique may then be described as follows: We take the first equation of (1) and divide by its leading coefficient, a_{11}, to get

TABLE I

Abbreviated Doolittle technique; forward solution

Eq.	Operation	Notation	x_1	x_2	x_3	x_4	
I		a_{i1}	a_{11}	a_{21}	a_{31}	a_{41}	a_{51}
II		a_{i2}	a_{12}	a_{22}	a_{32}	a_{42}	a_{52}
III		a_{i3}	a_{13}	a_{23}	a_{33}	a_{43}	a_{53}
IV		a_{i4}	a_{14}	a_{24}	a_{34}	a_{44}	a_{54}
V	I repeated	a_{i1}	a_{11}	a_{21}	a_{31}	a_{41}	a_{51}
VI	V divided by a_{11}	$b_{i1} = \dfrac{a_{i1}}{a_{11}}$	1	b_{21}	b_{31}	b_{41}	b_{51}
VII	II—b_{21} V	$a_{i2\cdot 1} = a_{i2} - a_{i1}b_{21}$	$a_{12\cdot 1}$	$a_{22\cdot 1}$	$a_{32\cdot 1}$	$a_{42\cdot 1}$	$a_{52\cdot 1}$
VIII	VII divided by $a_{22\cdot 1}$	$b_{i2\cdot 1} = \dfrac{a_{i2\cdot 1}}{a_{22\cdot 1}}$	$b_{12\cdot 1}$	1	$b_{32\cdot 1}$	$b_{42\cdot 1}$	$b_{52\cdot 1}$
IX	III—b_{31} (V)—$b_{32\cdot 1}$ VII	$a_{i3\cdot 12} = a_{i3} - a_{i1}b_{31} - a_{i2\cdot 1}b_{32\cdot 1}$	$a_{13\cdot 12}$	$a_{23\cdot 12}$	$a_{33\cdot 12}$	$a_{43\cdot 12}$	$a_{53\cdot 12}$
X	IX divided by $a_{33\cdot 12}$	$b_{i3\cdot 12} = \dfrac{a_{i3\cdot 12}}{a_{33\cdot 12}}$	$b_{13\cdot 12}$	$b_{23\cdot 12}$	1	$b_{43\cdot 12}$	$b_{53\cdot 12}$
XI	IV—b_{41} V—$b_{42\cdot 1}$ VII—$b_{43\cdot 12}$ IX	$a_{i4\cdot 123} = a_{i4} - a_{i1}b_{41} - a_{i2\cdot 1}b_{42\cdot 1} - a_{i3\cdot 12}b_{43\cdot 12}$	$a_{14\cdot 123}$	$a_{24\cdot 123}$	$a_{34\cdot 123}$	$a_{44\cdot 123}$	$a_{54\cdot 123}$
XII	XI divided by $a_{44\cdot 123}$	$b_{i4\cdot 123} = \dfrac{a_{i4\cdot 123}}{a_{44\cdot 123}}$	$b_{14\cdot 123}$	$b_{24\cdot 123}$	$b_{34\cdot 123}$	1	$b_{54\cdot 123}$

$$\sum_{i=1}^{n} b_{i1} x_i = b_{n+1,1}, \qquad \text{where} \quad b_{i1} = \frac{a_{i1}}{a_{11}}, \tag{2}$$

and we then form

$$\sum_{i=1}^{n} a_{i2\cdot 1} x_i = a_{n+1,2\cdot 1} \qquad \text{with} \quad a_{i2\cdot 1} = a_{i2} - a_{i1} b_{21}. \tag{3}$$

We then divide by $a_{22\cdot 1}$ and get

$$\sum_{i=1}^{n} b_{i2\cdot 1} x_i = b_{n+1,2\cdot 1} \qquad \text{with} \quad b_{i2\cdot 1} = \frac{a_{i2\cdot 1}}{a_{22\cdot 1}}. \tag{4}$$

We next form

$$\sum_{i=1}^{n} a_{i3\cdot 12} x_i = a_{n+1,3\cdot 12} \qquad \text{with} \quad a_{i3\cdot 12} = a_{i3} - a_{i1} b_{31} - a_{i2\cdot 1} b_{32\cdot 1}, \tag{5}$$

and

$$\sum_{i=1}^{n} b_{i3\cdot 12} x_i = b_{n+1,3\cdot 12} \qquad \text{with} \quad b_{i3\cdot 12} = \frac{a_{i3\cdot 12}}{a_{33\cdot 12}}. \tag{6}$$

This process is continued so that, in general, we have

$$\sum_{i=1}^{n} a_{ij\cdot 12\cdots j-1} x_i = a_{n+1,j\cdot 12\cdots j-1}, \qquad j = 1, 2, \cdots, n \tag{7}$$

and

$$\sum_{i=1}^{n} b_{ij\cdot 12\cdots j-1} x_i = b_{n+1,j\cdot 12\cdots j-1}, \qquad j = 1, 2, \cdots, n \tag{8}$$

with

$$\begin{aligned} a_{ij\cdot 12\cdots j-1} = a_{ij} - a_{i1} b_{j1} - a_{i2\cdot 1} b_{j2\cdot 1} - a_{i3\cdot 12} b_{j3\cdot 12} - \cdots \\ - a_{i,j-2\cdot 12\cdots j-3} b_{j,j-2\cdot 12\cdots j-3} - a_{i,j-1\cdot 12\cdots j-2} b_{j,j-1\cdot 12\cdots j-2} \end{aligned} \tag{9}$$

and

$$b_{ij\cdot 12\cdots j-1} = \frac{a_{ij\cdot 12\cdots j-1}}{a_{jj\cdot 12\cdots j-1}}. \tag{10}$$

It is to be noted that the n equations (1) are transformed by this process to the n auxiliary equations of (7) or (8). The solutions of (1) are also solutions of these auxiliary equations since the auxiliary equations are linear combinations of (1). It is our purpose to show that the prediagonal coefficients of these auxiliary equations are always 0 so that these auxiliary equations feature a decreasing number of variables.

We may use the term primary subscripts to indicate the first two subscripts and the term secondary subscripts to indicate the later subscripts which specify the order of elimination of the variables. The "order" of the coefficient is then equal to the number of secondary subscripts.

The formula (9) gives the matrix of the final Doolittle set of equations. At each stage of the reduction one can write down a formula for all the elements in the matrix at that stage. Thus one can write the coefficients of order h, $a_{ij\cdot 12\cdots h}$, in terms of coefficients of order less than h,

$$\begin{aligned} a_{ij\cdot 12\cdots h} = a_{ij} - a_{i1}b_{j1} - a_{i2\cdot 1}b_{j2\cdot 1} - \cdots \\ - a_{i,h-1\cdot 12\cdots h-2}b_{j,h-1\cdot 12\cdots h-2} - a_{ih\cdot 12\cdots h-1}b_{jh\cdot 12\cdots h-1}\,. \end{aligned} \tag{11}$$

It follows at once that

$$\begin{aligned} a_{ij\cdot 12\cdots h} &= a_{ij\cdot 12\cdots h-1} - a_{ih\cdot 12\cdots h-1}\,b_{jh\cdot 12\cdots h-1} \\ &= a_{ij\cdot 12\cdots h-1} - \frac{a_{ih\cdot 12\cdots h-1}\,a_{jh\cdot 12\cdots h-1}}{a_{hh\cdot 12\cdots h-1}}\,. \end{aligned} \tag{12}$$

4. Some theorems on the interchangeability of subscripts. Our main objective is to prove that the prediagonal terms are zero. In order to do this we first prove some theorems dealing with the primary and secondary subscripts.

THEOREM 1: *The value of $a_{ij\cdot\ldots h}$ is not changed if the primary subscripts are interchanged.* This theorem which might be stated "The matrix of the coefficients of a given order is symmetric" follows from the symmetry of the matrix of coefficients of zero order. We can show that the symmetry of the matrix having coefficient of order h follows at once from the symmetry of the matrix having coefficients of order $h - 1$ by comparing the value $a_{ij\cdot\ldots h}$ with that of $a_{ji\cdot\ldots h}$ obtained by dual substitution in (12). Since the matrix of zero order coefficients is symmetric by hypothesis, it follows that the matrices of the coefficients of order 1, 2, 3, 4, etc., are in turn symmetric.

THEOREM 2: *Any pair of consecutive secondary subscripts may be interchanged without changing the value of the coefficient.* This theorem indicates that, within prescribed limits, the order of elimination does not have any effect on the result.

Consider the coefficient $a_{ij\cdot\ldots kl\ldots}$ having r secondary subscripts before the k and s secondary subscripts after the l and consider the corresponding coefficient $a_{ij\cdot\ldots lk\ldots}$ which results from an interchange of k and l. These coefficients can be expressed by continued use of (12) in terms of coefficients of order $r + 2$. The resulting expansion of $a_{ij\cdot\ldots kl\ldots}$ is equivalent to that of $a_{ij\cdot\ldots lk\ldots}$ with the interchange of the l and the k. It follows that the theorem is true if $a_{ij\cdot\ldots lk} = a_{ij\cdot\ldots kl}$. Now a double application of (12) to $a_{ij\cdot\ldots lk}$ leads to the expansion in terms of coefficients of order r (using the notation $a_{ij\cdot}$ to indicate the coefficient of the r-th order)

$$a_{ij\cdot\ldots kl} = a_{ij\cdot} - \frac{a_{ik\cdot}\,a_{jk\cdot}}{a_{kk\cdot}} - \frac{\left(a_{il\cdot} - \dfrac{a_{ik\cdot}\,a_{lk\cdot}}{a_{kk\cdot}}\right)\left(a_{jl\cdot} - \dfrac{a_{jk\cdot}\,a_{lk\cdot}}{a_{kk\cdot}}\right)}{a_{ll\cdot} - \dfrac{a_{lk\cdot}^2}{a_{kk\cdot}}}\,. \tag{13}$$

Then $a_{ij\cdot\ldots lk}$ is expanded similarly, the difference is formed and found to be zero. It follows that the theorem is true.

The application of Theorem 2 with the continued interchange of successive secondary subscripts in all possible ways leads at once to

THEOREM 3: *The secondary subscripts may be interchanged in all possible ways without changing the value of the coefficient.* This theorem might be stated "The value of the resulting coefficient is independent of the order of elimination." This is the sort of result one would expect to find and indeed, some may feel that it is intuitively evident, but this formal proof is presented for those who desire a more rigorous approach.

Theorem 3 enables us to prove Theorem 4 which may be stated: *The value of $a_{ij.12\cdots n}$ is always zero if at least one of the secondary subscripts is equal to one of the primary subscripts.*

Suppose i is this subscript. Then by Theorem 3, i may be placed in the final position. Now by (12) we have

$$a_{ij.\cdots i} = a_{ij.\cdots} - \frac{a_{ij.\cdots}\, a_{ii.\cdots}}{a_{ii.\cdots}} = 0.$$

A similar statement holds if j appears among the secondary subscripts.

5. The vanishing of the prediagonal entries. As an application of Theorem 4 we can show that the prediagonal entries are identically zero and this is exactly what is needed to establish the validity of the forward Doolittle process. It is to be noted that the prediagonal entries are of form $a_{ij.12\cdots j-1}$ with $i < j$. Then i must equal one of the secondary subscripts and the term is zero.

It follows that no entries need be made to the left of the diagonal in the Abbreviated Doolittle solution and, indeed, no entries need be made in the original matrix below the main diagonal. A numerical problem is presented in the next section.

6. Illustration. The Abbreviated Doolittle technique is illustrated in Table II. This illustration is essentially an illustration of a previous article [2] and serves as the basis, in a later section, for expansion into the standard Doolittle solution. The check is shown in the right hand column and the back solution is indicated. The check entries for the first matrix are obtained by adding the entries in the row to the main diagonal and then adding the entries in the column. All other check entries are obtained by adding the entries in the row.

The solution is easily made once it is understood and results from continued application of formula (9). For example

$$a_{54\cdot 123} = a_{54} - a_{51}b_{41} - a_{52\cdot 1}b_{42\cdot 1} - a_{53\cdot 12}b_{43\cdot 12}$$

and this is

$$a_{54\cdot 123} = .8000 - (.2000)(.6000) - (.3200)(.1905) - (.4619)(-.1612) = .6935$$

(see the underscored entries of Table II). Terms of this sort are easily computed if a calculating machine, and especially so if one equipped with automatic

positive and negative multiplication, is available. The back solution too is easily accomplished with a machine. It is only necessary to substitute in turn in each of the "b" equations. Thus the value of x_1 is $\frac{a_{54\cdot123}}{a_{44\cdot123}} = b_{54\cdot123}$, the value of x_2 is $b_{53\cdot12} - b_{43\cdot12}b_{54\cdot123} = b_{53\cdot124}$, that of x_3 is $b_{52\cdot1} - b_{42\cdot1}b_{54\cdot123} - b_{32\cdot1}b_{53\cdot124} = b_{52\cdot134}$, etc. The back solution of the check is treated similarly.

7. A variation in technique. Before proceeding with the presentation of a standard Doolittle solution it seems wise to indicate another possible variation in the technique in addition to the division by the diagonal coefficient rather than its negative. It is possible to obtain the Doolittle solution by using the fixed entry from the first of the equivalent rows in place of using the fixed "b" entry and the variable "a". This results from the fact that

$$a_{ik\cdot\ldots}\, b_{jk\cdot\ldots} = a_{jk\cdot\ldots}\, b_{ik\cdot\ldots} \left(= \frac{a_{ik\cdot\ldots}\, a_{jk\cdot\ldots}}{a_{kk\cdot\ldots}}\right). \tag{14}$$

Thus in Table II the value $a_{54\cdot123}$ can be obtained with the use of

$$a_{54\cdot123} = a_{54} - a_{41}b_{51} - a_{42\cdot1}b_{52\cdot1} - a_{43\cdot12}b_{53\cdot12}$$

as readily as with the use of

$$a_{54\cdot123} = a_{54} - a_{51}b_{41} - a_{52\cdot1}b_{42\cdot1} - a_{53\cdot12}b_{43\cdot12}\,.$$

See the boxed entries of Table II.

There seems to be no real choice between these techniques. The fixed "b" is traditional in the standard Doolittle solution while the abbreviation of the method of single division leads to a fixed "a". The point to be emphasized here is that either the fixed "a" or the fixed "b" can be used. Also (14) is used in the next section in supplying details for the check portion of a standard Doolittle method.

8. The standard Doolittle method. If no computing machine is available or if a more detailed solution is desired, it is preferable to record the individual products of (9) and thus arrive at the standard Doolittle method. (The division by the diagonal coefficient rather than its negative is not a fundamental difference.) The standard Doolittle method, from this point of view, is an expanded form of the Abbreviated Doolittle method with more details added. Its validity then follows from the validity of the Abbreviated Doolittle method. While it is not true that all prediagonal terms vanish in the standard Doolittle method, and this fact complicates the check by row sums, yet the prediagonal $a_{ij\cdot\ldots}$ (and $b_{ij\cdot\ldots}$) are all zero.

The standard Doolittle method is presented in Table III. Some remarks should be made about the non-recorded terms, the two check solutions, and the back solution.

The blanks (—) indicate non zero entries which are usually not presented in a

Doolittle solution. They should be considered however if the first check method is to be used.

The first check method, which is the logical extension of the check method of the Abbreviated Doolittle solution, has been outlined by Ezekial [5]. The row sum is the sum of all the entries in the row whether recorded or not. In order to check, it is necessary to add these unrecorded entries, and they are available

TABLE II

Abbreviated Doolittle Solution; illustration

x_1	x_2	x_3	x_4		Check
1.0000	.4000	.5000	.6000	.2000	2.7000
—	1.0000	.3000	.4000	.4000	2.5000
—	—	1.0000	.2000	.6000	2.6000
—	—	—	1.0000	.8000	3.0000
1.0000	.40000	.5000	$\boxed{.6000}$	.2000	2.7000
1.0000	.40000	.5000	.6000	$\boxed{.2000}$	2.7000
	.8400	.1000	$\boxed{.1600}$	.3200	1.4200
	1.0000	.1190	.1905	$\boxed{.3810}$	1.6905
		.7381	$\boxed{-.1190}$	.4619	1.0810
		1.0000	−.1612	$\boxed{.6258}$	1.4646
			.5903	.6935	1.2837
			1.0000	1.1748	2.1747
		1.0000		.8152	1.8152
	1.0000			.0602	1.0602
1.0000				−.9366	.0635

in the columns above if we make use of formula (12). Thus, if we wish to check the value $\sum_{i=1}^{5} a_{i1}b_{41} = 1.6200$, we have

$$a_{11}b_{41} + a_{21}b_{41} + a_{31}b_{41} + a_{41}b_{41} + a_{51}b_{41} =$$

$$a_{41} + a_{41}b_{21} + a_{41}b_{31} + a_{41}b_{41} + a_{51}b_{41} =$$

$$.6000 + .2400 + .3000 + .3600 + .1200 = 1.6200.$$

Another check method, which is recommended by Peters and Van Voorhis [6] sums the entries in the row only over those columns which are to be recorded.

This is presented as check method 2 of Table III. As is to be expected, the check values of the a's and b's of the last two rows of each matrix are in agreement.

It might be noted that one may use the first check method without checking the intermediate steps (the sums for each row) if he checks the sums for the last two rows of each matrix.

TABLE III

Doolittle solution, with checks

Notation	x_1	x_2	x_3	x_4		Check Method 1	Check Method 2		
a_{i1}	1.0000	.4000	.5000	.6000	.2000	2.7000	2.7000		
a_{i2}	—	1.0000	.3000	.4000	.4000	2.5000	2.1000		
a_{i3}	—	—	1.0000	.2000	.6000	2.6000	1.8000		
a_{i4}	—	—	—	1.0000	.8000	3.0000	1.8000		
a_{i1}	1.0000	.4000	.5000	.6000	.2000	2.7000	2.7000		
b_{i1}	1.0000	.4000	.5000	.6000	.2000	2.7000	2.7000		
a_{i2}	—	1.0000	.3000	.4000	.4000	2.5000	2.1000		
$a_{i1}b_{21}$	—	.1600	.2000	.2400	.0800	1.0800	.6800		
$a_{i2\cdot1}$		.8400	.1000	.1600	.3200	1.4200	1.4200		
$b_{i2\cdot1}$		1.0000	.1190	.1905	.3810	1.6905	1.6905		
a_{i3}	—	—	1.0000	.2000	.6000	2.6000	1.8000		
$a_{i1}b_{31}$	—	—	.2500	.3000	.1000	1.3500	.6500		
$a_{i2\cdot1}b_{32\cdot1}$		—	.0119	.0190	.0381	.1690	.0690		
$a_{i3\cdot12}$			.7381	−.1190	.4619	1.0810	1.0810		
$b_{i3\cdot12}$			1.0000	−.1612	.6258	1.4646	1.4646		
a_{i4}	—	—	—	1.0000	.8000	3.0000	1.8000		
$a_{i1}b_{41}$	—	—	—	.3600	.1200	1.6200	.4800		
$a_{i2\cdot1}b_{42\cdot1}$		—	—	.0305	.0610	.2705	.0914		
$a_{i3\cdot12}b_{43\cdot12}$			—	.0192	−.0745	−.1743	−.0553		
$a_{i4\cdot123}$				.5903	.6935	1.2838	1.2839		
$b_{i4\cdot123}$				1.0000	1.1748	2.1748			
$b_{i3\cdot124}$			1.0000	−.1894	.8152	1.81532	−.3506		
$b_{i2\cdot134}$		1.0000	.0970	.2238	.0602	1.0602	.4143	.2160	
$b_{i1\cdot234}$	1.0000	.0241	.4076	.7049	−.9366	.0634	1.3049	.9076	.4241

The back solution is carried out as in Table II. If no computing machine is available or if the detailed steps are desired they may be indicated as in Table III. The entries in the box under the x_4 column are respectively $b_{54\cdot123}b_{43\cdot12}$, $b_{54\cdot123}b_{42\cdot1}$, and $b_{54\cdot123}b_{41}$. Those in the preceding column are $b_{53\cdot124}b_{32\cdot1}$ and $b_{53\cdot124}b_{31}$. The other entry is $b_{52\cdot134}b_{21}$. The values of the coefficients are obtained by subtracting these row entries from the constant term of the corresponding "b" equation. Thus, $b_{53\cdot124} = (.6258) - (-.1894)$; $b_{52\cdot134} =$

(.3810) − .0970 − .2238, etc. The back solution of check method 1 agrees with that of check method 2. A form for accomplishing the back solution of the check is indicated at the right. It is not necessary to complete the back solution of the check if it is not desired, and indeed, there are some who feel that the use of the row sum check is unnecessary with modern computing machines [7]. The basic check is substitution in the original equations.

9. Summary. The chief purpose of this paper is to show that the Doolittle technique actually leads to a set of equations featuring a decreasing number of unknowns. This is accomplished by the introduction of an appropriate notation to describe the process and the establishment of certain theorems which serve to validate the process. These theorems are of some interest aside from the application made here. It is a secondary purpose of this paper to emphasize the practicability and theoretical advantages (relative ease of calculating, theoretically more accurate, less chance for numerical error, less recording, less time consuming, more compact, and more easily checked) of the Abbreviated Doolittle method and to explain and illustrate possible variations in technique in the forward and check (by row sums) portions of the standard Doolittle solution. It should be noted that the notation suggested is very useful in providing an easy development of various theorems used in multiple and partial correlation studies, the presentation of which is not the purpose of the present paper.

REFERENCES

[1] M. H. Doolittle, "Method employed in the solution of normal equations and the adjustment of a triangulation," *U. S. Coast and Geodetic Survey Report* (1878), pp. 115–120.

[2] P. S. Dwyer, "The solution of simultaneous equations," *Psychometrika*, Vol. 6 (1941), pp. 101–129.

[3] A. C. Aitken, "Studies in practical mathematics I. The evaluation, with applications, of a certain triple product matrix," *Roy. Soc. Edin. Proc.*, Vol. 57 (1937), pp. 172–181.

[4] C. F. Gauss, "Supplementum theoriae combinationis observationum erroribus minimis obnoxiae," *Werke*, Vol. 4 (1873), pp. 69–71.

[5] Mordecai Ezekial, *Methods of Correlation Analysis*. John Wiley and Sons, Inc., New York (1930), pp. 362–364.

[6] C. C. Peters, and W. R. Van Voorhis, *Statistical Procedures and Their Mathematical Bases*, McGraw Hill (1940), pp. 228–229.

[7] A. K. Kurtz, "The use of the Doolittle method in obtaining related multiple correlation coefficients," *Psychometrika*, Vol. 1 (1936), pp. 45–51.

NOTES

This section is devoted to brief research and expository articles, notes on methodology and other short items.

A PROBLEM IN ESTIMATION

By Joseph F. Daly

The Catholic University of America

Several recent psychological studies in the field of memory testing [1], [2], [3] have suggested the following problem. Let each individual E in our population be characterized by the variates $y^1, \cdots, y^p; y^{p+1}, \cdots, y^{p+t}$ $(p > t)$. Suppose, however, that circumstances make it impossible for us to observe the last t variates. For example, we may think of $y^1, \cdots, y^p$ as an individual's scores on a battery of tests, and think of $y^{p+1}, \cdots, y^{p+t}$ as measures of certain psychological characteristics which, though affecting the individual's performance, are not subject to direct observation. To make up for this, assume that we have a theory which tells us that if $y^{p+1}, \cdots, y^{p+t}$ are held constant, then the observable y's are dependent upon them according to a specified regression equation

$$y^i = x^i_\mu y^\mu, \qquad (i = 1, \cdots, p; \mu = p + 1, \cdots, p + t).$$

Somewhat more precisely, we assume the distribution laws

$$(1) \quad f(y^1, \cdots, y^{p+t}) = (2\pi)^{-\frac{1}{2}(p+t)} \mid A_{rs} \mid^{\frac{1}{2}} \exp\left\{-\tfrac{1}{2}A_{rs}(y^r - a^r)(y^s - a^s)\right\},$$

(where $r, s = 1, \cdots, p + t$, and repeated indices are to be summed according to the usual convention) and

$$(2) \quad f(y^1, \cdots, y^p \mid y^{p+1}, \cdots, y^{p+t}) = (2\pi\sigma^2)^{-\frac{1}{2}p} \exp\left\{-\frac{1}{2\sigma^2}\sum_i (y^i - x^i_\mu y^\mu)^2\right\}.$$

The x^i_μ are supposed to be known, but except for the conditions imposed by (1) and (2) nothing is known about the quantities A_{rs}, a^r, and σ^2. Having observed the test scores y^i_α $(\alpha = 1, \cdots, N)$ obtained by N individuals E_α drawn at random from the population, we wish to estimate the values $y^{p+1}_\alpha, \cdots, y^{p+t}_\alpha$ corresponding to each E_α, and the essential parameters in the distribution law (1), particularly the variances and covariances of $y^{p+1}, \cdots, y^{p+t}$.

We can easily find optimum estimates of the y^μ_α by applying the method of maximum likelihood to the function (2) after substituting for the y^i the scores y^i_α obtained by the individual in question. Thus if we write

$$v_{\mu\nu} = x_\mu^i x_\nu^i, \qquad \| v^{\mu\nu} \| = \| v_{\mu\nu} \|^{-1},$$

(assuming thereby that the rank of the matrix $\| x_\mu^i \|$ is t) we have

$$\hat{y}_\alpha^\mu = v^{\mu\nu} x_\nu^i y_\alpha^i. \tag{3}$$

These estimates are unbiased in the sense that the expected value of $\hat{y}^\mu$ calculated from the distribution law (2) is y^μ.

But when we come to estimate the variances and covariances involved in (1), the procedure is less straightforward. Under the present circumstances we cannot use the expression

$$\frac{1}{N-1} \sum_\alpha (y_\alpha^\mu - \bar{y}^\mu)(y_\alpha^\nu - \bar{y}^\nu), \tag{4}$$

for the sample covariance of y^μ and y^ν. We might, of course, try substituting the estimates $\hat{y}_\alpha^\mu$ from (3) for the unknown y_α^μ in (4). But this expedient will in general produce a biased estimate. Denoting the required covariance by $A^{\mu\nu}$ (the element in the appropriate position in the inverse of the matrix $\| A_{rs} \|$), we find as a matter of fact that the expected value of (4) when the y_α^μ are replaced by their estimates $\hat{y}_\alpha^\mu$ is

$$A^{\mu\nu} + \sigma^2 v^{\mu\nu}. \tag{5}$$

This bias may or may not be important in any given case. But it can conceivably be quite serious if the $A^{\mu\nu}$ are relatively small, especially if such expressions are employed in the usual way to estimate the correlation coefficient rather than the covariance.

Perhaps the most logical way to attack the problem is through the joint distribution of $y^1, \cdots, y^p$ alone, obtainable by integrating the undesirable variates $y^{p+1}, \cdots, y^{p+t}$ out of (1). We therefore consider

$$f(y^1, \cdots, y^p) = (2\pi)^{-\frac{1}{2}p} \,|\, \tilde{A}_{ij} \,|^{\frac{1}{2}} \exp\{-\tfrac{1}{2}\tilde{A}_{ij}(y^i - a^i)(y^j - a^j)\}, \tag{6}$$

where

$$\tilde{A}_{ij} = A_{ij} - A_{i\mu}B^{\mu\nu}A_{\nu j}, \qquad \| B^{\mu\nu} \| = \| A_{\mu\nu} \|^{-1}.$$

Moreover, when account is taken of (2), we find that we must have

$$A_{ij} = \frac{\delta_{ij}}{\sigma^2}, \qquad A_{i\mu} = -\frac{x_\mu^i}{\sigma^2} \qquad a^i = x_\mu^i a^\mu$$

(δ_{ij} being Kronecker's delta). If we now form the likelihood function $\prod_{\alpha=1}^{N} f(y_\alpha^1, \cdots, y_\alpha^p)$ from (6) for our sample, and set its derivatives with respect to the a^μ, σ^2, and the $B^{\mu\nu}$, equal to zero, we arrive, after some simplification, at the equations

$$a^\mu = v^{\mu\nu} x_\nu^i \bar{y}^i = \frac{1}{N} \sum \hat{y}_\alpha^\mu, \qquad \text{[cf. (3)]}$$

$$
(7) \quad \begin{aligned}
&\left\{A^{ij} - \frac{1}{N}\sum_{\alpha}(y^i_\alpha - x^i_\mu a^\mu)(y^j_\alpha - x^j_\nu a^\nu)\right\}\delta_{ij} = 0,\\
&\left\{A^{ij} - \frac{1}{N}\sum_{\alpha}(y^i_\alpha - x^i_\mu a^\mu)(y^j_\alpha - x^j_\nu a^\nu)\right\}x^i_\sigma x^j_\tau = 0,\\
&A^{ij} = \sigma^2\delta^{ij} + x^i_\mu A^{\mu\nu}x^j_\nu,
\end{aligned}
$$

for determining the maximum likelihood estimates. The first of equations (7) is already solved for the a^μ, and the solution of the simultaneous equations for the remaining essential parameters yields the estimates

$$
(8) \quad \hat{\sigma}^2 = \frac{1}{N(p-t)}\sum_{\alpha,i}(y^i_\alpha - x^i_\mu \hat{y}^\mu_\alpha)^2
$$

$$
(9) \quad \hat{A}^{\mu\nu} = \frac{1}{N}\sum_{\alpha}(\hat{y}^\mu_\alpha - \hat{a}^\mu)(\hat{y}^\nu_\alpha - \hat{a}^\nu) - v^{\mu\nu}\hat{\sigma}^2.
$$

A considerable amount of algebraic manipulation is required to put the solutions in the form given above; but since the results are about what one would expect in view of (5), we omit the details. As is often the case, some bias remains in the "optimum" estimates (9). However, this can be eliminated by writing $N - 1$ in place of N. The estimate (8) of σ^2 is unbiased as it stands.

REFERENCES

[1] T. V. Moore, "The analysis of association by its equational constants," *Aspects of the New Scholastic Philosophy*, New York, Benziger Brothers, 1932, pp. 181–225.

[2] Regis Holland, "The development of logical and rote memory," *Studies in Psychology and Psychiatry*, the Catholic University of America, Vol. 4, no. 8 (1940).

[3] Helen deSales Forrest, "Correlations between the constants in the curve of learning," *Studies in Psychology and Psychiatry*, the Catholic University of America, Vol. 5, no. 1 (1941).

CONFIDENCE LIMITS FOR AN UNKNOWN DISTRIBUTION FUNCTION

By A. Kolmogoroff

Moscow, U.S.S.R.

Let $x_1, x_2, \cdots, x_n$ be mutually independent random variables following the same distribution law

$$
(1) \quad P\{x_i \leq \xi\} = F(\xi).
$$

A recent paper by A. Wald and J. Wolfowitz[1] deals with the problem of using

---

[1] A. Wald and J. Wolfowitz, "Confidence limits for distribution functions," *Annals of Math. Stat.*, Vol. 10 (1939), pp. 105–118.

the observable values of the x's to estimate the function $F(\xi)$. In this connection it may be useful to recall the following results published by me in 1933.[2]

Put

$$F_n(\xi) = \frac{N(\xi)}{n} \tag{2}$$

where $N(\xi)$ denotes the number of those x's whose observed values do not exceed ξ.

THEOREM 1: *If the function $F(\xi)$ is continuous then the distribution law of the quantities*

$$D_n = \sup | F(\xi) - F_n(\xi) | \sqrt{n} \tag{3}$$

does not depend on $F(\xi)$.

Denote by $\Phi_n(\lambda)$ the value of the probability $P\{D_n \leq \lambda\}$ which is common to all continuous distribution functions $F(\xi)$.

THEOREM 2: *For n tending to infinity, the distribution function $\Phi_n(\lambda)$ tends to*

$$\Phi(\lambda) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2\lambda^2} \tag{4}$$

uniformly with respect to λ.

A more elementary proof of Theorem 2 was given by N. Smirnoff in 1939.[3] Another paper by the same author[4] gives a table of the function $\Phi(\lambda)$.

Without the assumption that $F(\xi)$ is continuous, we easily obtain

THEOREM 3: *Whatever be the distribution function $F(\xi)$,*

$$P\{D_n \leq \lambda\} \geq \Phi_n(\lambda). \tag{5}$$

Theorems 1 and 3 giving the exact lower bound of the probability that $F_n(\xi)$ will satisfy the inequality

$$| F(\xi) - F_n(\xi) | \leq \frac{\lambda}{\sqrt{n}} \tag{6}$$

for all values of ξ, can be used to establish confidence limits for $F(\xi)$ corresponding to the confidence coefficient

$$\alpha = \Phi_n(\lambda). \tag{7}$$

These confidence limits will be free from any restriction concerning the nature of the function $F(\xi)$.

---

[2] A. Kolmogoroff, "Sulla determinatione empirica di una legge di distributione," *Giornale dell'Istituto Italiano degli Attuari*, Vol. 4 (1933), pp. 83-91.

[3] N. Smirnoff, "Sur les écarts de la courbe de distribution empirique," *Recueil Math. de Moscou*, Vol. 6 (1939), pp. 3-26.

[4] N. Smirnoff, "On the estimation of the discrepancy between empirical curves of distribution for two independent samples," *Bulletin de l'Université de Moscou, Série internationale (Mathématiques)*, Vol. 2, fasc. 2 (1939).

For sufficiently large values of n we can use the limiting distribution (4) and write

$$\alpha = \Phi(\lambda). \tag{8}$$

The following short table, based on that of Smirnoff,[4] gives the values of λ corresponding to a few chosen confidence coefficients α.

TABLE OF λ

α	λ
.95	1.35
.98	1.52
.99	1.63
.995	1.73
.998	1.86
.999	1.95

Smirnoff's paper[4] contains still another application of the function $\Phi(\lambda)$. Denote by $x_1', x_2', \cdots, x_{n_1}'$ and $x_1'', x_2'', \cdots, x_{n_2}''$ two sequences of mutually independent random variables following the same probability law $F(\xi)$. Let further $F_{n_1}(\xi)$ and $F_{n_2}(\xi)$ be two random step functions corresponding to these series, defined as in (2). Smirnoff proves then the following

THEOREM 4: *If the probability law $F(\xi)$ is continuous, then the probability*

$$P\left\{\sup | F_{n_1}(\xi) - F_{n_2}(\xi) | \leq \lambda \sqrt{\frac{n_1 + n_2}{n_1 n_2}}\right\} = \Phi_{n_1, n_2}(\lambda) \tag{9}$$

is independent of the function $F(\xi)$. If n_1 and n_2 are indefinitely increased subject to the restriction that the ratio n_1/n_2 remains between two fixed numbers a_1 and a_2

$$0 < a_1 \leq \frac{n_1}{n_2} \leq a_2 < +\infty \tag{10}$$

then

$$\Phi_{n_1, n_2}(\lambda) \to \Phi(\lambda). \tag{11}$$

In the general case, where the probability law $F(\xi)$ is absolutely arbitrary we have

$$P\left\{\sup | F_{n_1}(\xi) - F_{n_2}(\xi) | \leq \lambda \sqrt{\frac{n_1 + n_2}{n_1 n_2}}\right\} \leq \Phi_{n_1, n_2}(\lambda). \tag{12}$$

Owing to the above results the quantity

$$D_{n_1, n_2} = \sup | F_{n_1}(\xi) - F_{n_2}(\xi) | \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \tag{13}$$

could be used as a criterion to test the hypothesis that the probability laws of the two series of observable variables are actually the same.

CORRECTIONS TO A PAPER ON THE UNIQUENESS PROBLEM OF MOMENTS

By M. G. Kendall

London, England

I wish to make certain corrections in my paper on "Conditions for Uniqueness in the Problem of Moments" (*Annals of Math. Stat.*, Vol. 11 (1940), p. 402). I thought I had succeeded in improving on results given earlier by Stieltjes, Lévy and Carleman, but this is not so.

Theorem 1 of the paper stated that a set of moments determines a distribution uniquely if $\sum_{r=0}^{\infty} \frac{\nu_r t^r}{r!}$ converges for some real non-zero t, ν_r being the absolute moment of order r. This is true, and a similar result has been proved by Lévy, but my proof contained a small lacuna. It was shown that the characteristic function $\phi(t)$ has a Taylor expansion which, under the conditions of the theorem, is convergent; but it has also to be shown that it is equal to the sum of that expansion. This may be seen as follows:

We have

$$e^{itx} - \sum_{r=0}^{n-1} \frac{(itx)^r}{r!} \leq \theta \frac{|tx|^n}{n!}, \qquad |\theta| \leq 1,$$

and hence, on taking mean values,

$$\left| \phi(t) - \sum_{r=0}^{n-1} \frac{(it)^r \mu_r}{r!} \right| \leq \frac{\nu_n t^n}{n!}.$$

Since by hypothesis $\frac{\nu_n t^n}{n!} \to 0$, $\phi(t)$ must be equal to the sum of its (convergent) Taylor expansion.

The principal error was a statement that $\nu_n^{1/n}/n$ must either tend to a limit or diverge. For this reason, the second theorem should run: a distribution determines a distribution uniquely if $\overline{\lim}\, \nu_n^{1/n}/n$ is finite (not $\lim \nu_n^{1/n}/n$ as originally stated). Theorem 3 should also be restated with the upper limit substituted for the limit therein.

Theorem 4 stated that a set of moments uniquely determines a distribution if $\Sigma \frac{1}{\nu_n^{1/n}}$ diverges. A rigorous proof is as follows:

The characteristic function obeys the relation

$$|\phi^{(n)}(t)| \leq \nu_n, \qquad n > 1$$

provided, of course, that ν_n exists. A theorem of Denjoy[1] states that if a function $f(x)$, defined in the segment (a, b), possesses derivatives of all orders therein,

---

[1]Arnaud Denjoy, "Sur les fonctions quasi-analytiques de variable réelle," *Comptes Rendus* Vol. 173 (1921), p. 1399.

if M_n is the maximum of $|f^{(n)}(x)|$ in the segment and if $\Sigma \frac{1}{M_n^{1/n}}$ is divergent, then $f(x)$ is completely determined by its value and that of its derivatives at a single point. $\phi(t)$ obeys the conditions of the theorem and by taking the point to be $t = 0$, theorem 4 follows.

I hope that this note will correct any misunderstandings that may have arisen on the main paper, and I regret that a number of circumstances, not the least of which is war, have made it impossible to forward the correction at an earlier date.

ANNOUNCEMENT CONCERNING COMPUTATION OF MATHEMATICAL TABLES

In the December, 1939, issue of the *Annals of Mathematical Statistics*, p. 399, there appeared an Announcement of the *Mathematical Tables Project*. This project is operated by the Work Projects Administration of New York City, as O. P. No. 265-2-97-11 under the technical supervision of Dr. A. N. Lowan. It is sponsored by the National Bureau of Standards, Dr. Lyman J. Briggs, Director.

In order to keep the readers of the *Annals* up-to-date on the progress of the work of the Project, information will be released from time to time.

The following list shows the status of work, as of October, 1941. The reader is referred to the December, 1939 issue of the *Annals* with respect to which n will denote the n^{th} item of Tables Published, Pn will denote the n^{th} item of Tables in Progress and Cn will denote the n^{th} item of Tables under Consideration.

Tables published. 1, 2, 3, *P*1, *P*2, *P*3, *P*4, *P*6(b), *P*6(c), *P*6(d), *P*6(e), *P*7, *C*7 and also

1. Table of Five-Point Lagrangian Interpolants for arguments ranging between 0 and 2 at intervals of 0.001.
2. Tables of Grid Coordinates (American Polyconic Projection) at 5 minute intervals of latitude and longitude for latitude from 70°N to 28°N and for latitude from 49°N to 72°N.
3. Table for Map Projections of Northwestern Extension of U. S.

Tables in process of reproduction. *P*5, *P*6(a), *P*8 and *C*1 for [0 (.001) 7 (.01) 50 (.1) 300 (1) 2,000 (10) 10,000; 12*D*] also

1. Tables of Section Moduli and Moments of Inertia for Structural Members used in Naval Architecture. (For the Bureau of Marine Inspection and Navigation.)
2. Tables of $Si(x)$ and $Ci(x)$ for x ranging from 10 to 100 at intervals of 0.001.

3. The zeros of the Legendre Polynomials up to the 16th order to 15 decimal places and the Weight Coefficients for Gauss' Mechanical Quadrature Formula.

Tables for which manuscripts are completed. $P9$, $P11$, $C6$, (the function x^y, instead of $A(x, y)$, has been tabulated to 15 places), and also

1. Table of $\int_0^x J_0(t)\, dt$ from 0 to 10 at intervals of 0.01 to 10 places.

Tables for which computations are completed. $P10$ (also tanh x, coth x), $C2$, $C3$, (change to $n = -21, -20 \cdots 0$) and also

1. Various hydraulic tables based on Kutter's and Manning's formulae. (Tabulation suggested by the War Department.)

2. Table of reciprocals of the integers from 100,000 to 200,000.

3. Table of the Associated Legendre Functions $P_n^m(x)$ and $Q_n^m(x)$ for n ranging between 1 and 10, and m between 0 and 4; for arguments x and ix where x ranges between 0 and 10 at intervals of 0.1. Also corresponding values for half-integral values of n and values of the functions for arguments in degrees. (Tabulation suggested by National Defense Research Committee.)

4. Tables of $R \sin \theta$ and $R \cos \theta$. $R = 1000\ (10)\ 10{,}000$, $\theta = 5(5)800$ (in mils).

Tables for which computations are in progress. $C3$ (for $n = 1, 2, \cdots 20$) and also

1. Table of the Bessel Functions $Y_0(z)$ and $Y_1(z)$ for the same complex arguments as in $J_0(z)$ and $J_1(z)$, mentioned in $P9$.

2. Tables of Length of Meridional Arc at one-minute intervals.

3. Tables of the Confluent Hypergeometric Function for selected values of the parameters.

4. Tables of three-point, four-point, six-point and seven-point Lagrangian Interpolants.

5. Table of Tchebysheff Polynomials.

Tables under consideration. $C4$ and also

1. Table of the first 10 powers of the reciprocals of the integers from 1 to 1,000.

2. Extensive tables of Elliptic Functions for both real and imaginary arguments.

3. A 12-place table of Inverse Circular and Hyperbolic Functions other than Arc tan x.

4. Table of the Integral $\int_0^x Y_0(t)\, dt$.

5. Tables of the non-periodic solutions of the Mathieu Differential Equation.

6. Table of the Error Functions for complex arguments (suggested by Federal Communication Commission).

7. Tables of the Unit-Sigma Functions and their integrals.

8. Tables of Circular Functions for Complex Arguments.

9. Tables of the Zeros of the Hermite and Laguerre Polynomials and of the corresponding Weight Factors in Gauss' Mechanical Quadrature Formula.

10. Table of Lamé Polynomials.

11. Table of Military Grid Coordinates for certain "Control Stations." (For the War Department.)

12. Tables of the Chi-Square Distribution and "Student's" t-distribution.

13. Tabulation of Fisher's A-, B-, and C- Distributions of the Multiple Correlation Coefficients.

The Project would welcome suggestions for the computation of new tables of interest in pure and applied mathematics, as well as information regarding computational work in progress elsewhere.

Communications should be addressed to Major Irving V. Huie, Administrator, Work Projects Administration, 70 Columbus Avenue, New York City.

Requests for copies of published tables should be addressed to Dr. Lyman J. Briggs, Director of the National Bureau of Standards, Washington, D. C.

REPORT OF THE CHICAGO MEETING OF THE INSTITUTE

The Fourth Summer Meeting of the Institute of Mathematical Statistics was held at The University of Chicago, Tuesday to Thursday, September 2 to 4, 1941, in conjunction with the meetings of the American Mathematical Society, the Mathematical Association of America, and the Econometric Society. The following sixty-eight members of the Institute attended the meeting:

R. L. Anderson, T. W. Anderson, K. J. Arnold, H. M. Bacon, Walter Bartky, W. D. Baten, A. A. Bennett, Paul Boschan, I. W. Burr, J. H. Bushey, W. E. Cederberg, W. G. Cochran, A. T. Craig, C. C. Craig, J. H. Curtiss, J. F. Daly, W. E. Deming, J. L. Doob, P. L. Dressel, P. S. Dwyer, Churchill Eisenhart, M. L. Elveback, H. P. Evans, C. H. Fischer, W. C. Flaherty, R. M. Foster, C. H. Graves, Louis Guttman, W. L. Hart, F. C. Hinds, A. S. Householder, E. V. Huntington, William Hurwitz, M. H. Ingraham, Dunham Jackson, Leo Katz, J. F. Kenney, L. A. Knowler, L. F. Knudsen, Tjalling Koopmans, C. F. Kossack, O. E. Lancaster, D. H. Leavens, B. A. Lengyel, W. G. Madow, J. N. Michie, A. M. Mood, J. E. Morton, Leah Naugle, Harold Nisselson, J. I. Northam, E. G. Olds, Oystein Ore, C. K. Payne, G. A. D. Preinreich, Francis Regan, Selby Robinson, C. F. Roos, M. M. Sandomire, Max Sasuly, Henry Scheffe, H. M. Schwartz, Harry Siller, J. H. Smith, M. E. Wescott, S. S. Wilks, E. W. Wilson, Gale Young.

The opening session, on Tuesday morning, was devoted to contributed papers on *Probability and Statistics* and was held jointly with the American Mathematical Society and the Econometric Society. The Chairman was Professor A. T. Craig, University of Iowa, and the following papers were presented:

1. *A geometric derivation of Fisher's z-transformation.*
   J. B. Coleman, University of South Carolina.
2. *Large sample distribution of the likelihood ratio.*
   Abraham Wald, Columbia University.
3. *On the integral equation of renewal theory.*
   (Read by title.)
   Willy Feller, Brown University.
4. *Cumulative frequency functions.*
   Irving Burr, Purdue University.
5. *On spherical probability distributions.*
   K. J. Arnold, Massachusetts Institute of Technology.
6. *Some observations on analysis of variance theory.*
   (Read by title.)
   Hilda Geiringer, Bryn Mawr College.
7. *On the asymptotic distribution of medians of samples from a multivariate population.*
   A. M. Mood, University of Texas.
8. *A problem of estimation.*
   J. F. Daly, Catholic University.

Abstracts of these papers follow this report.

On Tuesday afternoon a session was held jointly with the Econometric Society on *Time Series Analysis*. Under the chairmanship of Professor C. C. Craig of the University of Michigan, the following papers were presented:

1. *Is sampling theory applicable to economic time series?*
   Tjalling Koopmans, Penn Mutual Life Insurance Co., Philadelphia.
2. *Serial correlation.*
   R. L. Anderson, North Carolina State College.

The morning session on Wednesday was held jointly with the Econometric Society on *Curve Fitting*. The chair was held by Dr. J. Marschak of the New School for Social Research and the following papers were presented:

1. *Weights to compensate for transformation in curve fitting.*
   T. O. Yntema, University of Chicago and Cowles Commission.
2. *Curve fitting by cumulative addition.*
   John H. Smith, University of Chicago and Cowles Commission.

On Wednesday afternoon, Professor S. S. Wilks of Princeton University acted as chairman of a session on *Multivariate Analysis*. The following papers were read:

1. *On testing sets of means and discriminant analysis.*
   Abraham Wald, Columbia University.
2. *On tests of hypotheses concerning variances and covariances.*
   William G. Madow, Bureau of the Census.

The Josiah Willard Gibbs Lecture of the American Mathematical Society was delivered on Wednesday evening by Professor Sewall Wright of the University of Chicago. His topic was *Statistical Genetics and Evolution*.

On Thursday morning a joint session on *Demand and Supply Analysis* was held with the Econometric Society. At this session Dr. C. F. Roos of the Institute of Applied Econometrics presided, and the following papers were presented:

1. *Demand analysis for certain commodities based on income and budget data.*
   J. Marschak, New School for Social Research, and George Garvey, National Bureau of Economic Research.
2. *Derivation of elasticities of demand and supply: A direct method.*
   Oscar Lange, University of Chicago and Cowles Commission.
3. *On the workings of a general equilibrium system.*
   J. L. Mosak, University of Chicago and Cowles Commission.

An informal reception was held on Monday evening in the Judson Court Lounge. On Tuesday and Wednesday afternoons the ladies of the Mathematics Department of the University of Chicago served tea in the Eckhart Hall Common Room. After the joint session on Tuesday afternoon, the Cowles Commission for Research in Economics gave a tea in the Common Room of the Science Building. On Thursday evening a joint dinner of the four mathematical organizations was held in Hutchinson Commons, preceded by an informal reception at the Reynolds Club.

EDWIN G. OLDS,
Secretary

ABSTRACTS OF PAPERS

(Presented on September 2, 1941, at the Chicago Meeting of the Institute)

A Geometric Derivation of Fisher's z-transformation. J. B. COLEMAN, University of South Carolina.

In fitting points in a plane by a line so that the sum of the squares of the perpendicular deviations shall be a minimum, a second line is found for which the sum of the squares of the deviations is a maximum. Let Σd^2 be the sum of the squares of the deviations of the points from the minimum line, and ΣD^2 be the sum of the squares from the maximum line. Then $\Sigma D^2/\Sigma d^2 = (1 + r)/(1 - r)$. $\frac{1}{2} \log (1 + r)/(1 - r)$ is Fisher's z-transformation for testing the coefficient of correlation.

Large Sample Distribution of the Likelihood Ratio. ABRAHAM WALD, Columbia University.

The large sample distribution of the likelihood ratio has been derived by S. S. Wilks (*Annals of Math. Stat.*, Vol. 9 (1938)) in case of a linear composite hypothesis and under the assumption that the hypothesis to be tested is true. Here a general composite hypothesis is considered and the distribution in question is derived also in case that the hypothesis to be tested is not true. Let $f(x_1, \cdots, x_p, \theta_1, \cdots, \theta_k)$ be the joint probability density function of the variates $x_1, \cdots, x_p$ involving k unknown parameters $\theta_1, \cdots, \theta_k$. Denote by H_ω the hypothesis that the true parameter point $\theta = (\theta_1, \cdots, \theta_k)$ satisfies the equations $\xi_1(\theta) = \cdots = \xi_r(\theta) = 0$, $(r \leq k)$. Denote by λ_n the likelihood ratio statistic for testing H_ω on the basis of n independent observations on $x_1, \cdots, x_p$. For any parameter point θ let $\xi_{ij}(\theta) = \dfrac{\partial \xi_i(\theta)}{\partial \theta_j}$ and let $c_{ij}(\theta)$ be the expected value of $\dfrac{\partial \log f(x_1, \cdots, x_p, \theta)}{\partial \theta_i}$. $\dfrac{\partial \log f(x_1, \cdots, x_p, \theta)}{\partial \theta_j}$ calculated under the assumption that θ is the true parameter point. For any θ denote by $A(\theta)$ the matrix $\| \xi_{ij}(\theta) \|$ $(i = 1, \cdots, r; j = 1, \cdots, k)$ and let $\| \sigma_{ij}(\theta) \| = \| c_{ij}(\theta) \|^{-1}$, $(i, j = 1, \cdots, k)$. Let furthermore $\| \sigma^*_{uv}(\theta) \|$, $(u, v = 1, \cdots, r)$ be the matrix equal to the product $A(\theta) \cdot \| \sigma_{ij}(\theta) \| \cdot \bar{A}(\theta)$, where $\bar{A}(\theta)$ is the transpose of $A(\theta)$. Finally let $\| \bar{c}^*_{uv}(\theta) \| = \| \sigma^*_{uv}(\theta) \|^{-1}$, $(u, v = 1, \cdots, r)$. For each n and θ denote by $y_{1n}(\theta), \cdots, y_{rn}(\theta)$ a set of r variates which have a joint normal distribution with mean values $\sqrt{n}\xi_1(\theta), \cdots, \sqrt{n}\xi_r(\theta)$ and covariance matrix $\| \sigma^*_{uv}(\theta) \|$, $(u, v = 1, \cdots, r)$. Denote the quadratic form $\sum_{v=1}^{r} \sum_{u=1}^{r} y_{un}(\theta) y_{vn}(\theta) \bar{c}^*_{uv}(\theta)$ by $Q_n(\theta)$. It has been shown that under certain assumptions on $f(x_1, \cdots, x_p, \theta)$, $\xi_1(\theta), \cdots, \xi_r(\theta)$ we have $\lim_{n=\infty} \{P(-2 \log \lambda_n < t \mid \theta) - P[Q_n(\theta) < t \mid \theta]\} = 0$ uniformly in t and θ, where for any z $P(z < t \mid \theta)$ denotes the probability that $z < t$ holds under the assumption that θ is the true parameter point. The distribution of $Q_n(\theta)$ is known and has been treated in the literature. If H_ω is true, then $\xi_1(\theta) = \cdots = \xi_r(\theta) = 0$, and $Q_n(\theta)$ has the χ^2 distribution with r degrees of freedom.

On the Integral Equation of Renewal Theory. W. FELLER, Brown University.

As is well-known, the equation $U(t) = G(t) + \int_0^t U(t - x)\, dF(x)$ has frequently been

discussed, under different forms, in connection with the population theory, the theory of industrial replacement, etc. In the present paper it is shown that, using Tauberian theorems for Laplace integrals, it becomes possible to analyze in detail the asymptotic behavior of $U(t)$ as $t \to \infty$ and also to solve some other problems which have been discussed in the literature. Strict conditions for the validity of different methods to treat the equation are given together with some modifications found to be necessary. The paper will appear in the *Annals of Mathematical Statistics*.

Cumulative Frequency Functions. I. W. BURR, Purdue University.

Frequency and probability functions play a fundamental role in statistical theory and practice. They are, however, often inconvenient and difficult to use, since it is necessary to integrate or sum to find the probability for a given range. Theoretically the cumulative or integral frequency function would seem to be better adapted to determining such probabilities, since the latter can be found simply by a subtraction. The aim of this paper is to make a contribution toward the direct use of cumulative frequency functions. Some general properties and theory of cumulative functions are presented with particular emphasis upon certain moment functions adapted to such direct use. Both continuous and discrete cases are included. A list of possible cumulative functions is given and a particular one, $F(x) = 1 - (1 + x^c)^{-k-1}$, discussed fully. This function has properties which make it practicable and adaptable to a wide variety of distribution types. It well illustrates the possibilities of the cumulative approach.

On Spherical Probability Distributions. KENNETH J. ARNOLD, Massachusetts Institute of Technology.

Two methods of correspondence for circular distributions to the normal error function have led to non-constant absolutely continuous functions [See F. Zernike's article in *Handbuch der Physik* Vol. 3, pp. 477–478]. The corresponding distributions for the sphere are found. The case of diametrical symmetry for both circle and sphere is discussed. Tables of the probability integrals involved are given and an application in geology is included.

Some Observations on Analysis of Variance Theory. HILDA GEIRINGER, Bryn Mawr College.

The test functions used in analysis of variance present themselves in different classes of important problems. Their distribution has been determined and tabulated by *R. A. Fisher*[1] under the hypothesis that the chance variables are all *independent* of each other and subject to the *same normal* law. Consequently we can in this way test only the hypothesis that the theoretical populations have all these properties.

If it is not possible to determine the exact distribution of test functions under sufficiently general assumptions regarding the populations we may: (a) find an asymptotic solution of the problem, i.e. determine the distribution of the test functions *for large samples*.[2] Or (b) determine at least the mathematical expectations and the variances of the test functions for *appropriately general* populations and for *small samples*.

It is well known that the expectations of the two quadratic forms which are basic in the analysis of variance are *equal*, even if the n populations are not normal but equal to each other (Bernoulli series). But, in addition, we can prove the mathematical theorem that, under the same conditions the *expectation of their quotient equals one*. The next step consists in studying the case that the n distributions are not equal to each other and to investigate certain *inequalities* characteristic for the Lexis Series and Poisson Series. These different criteria are completed by the *computation of the variances* of the test functions.

---

[1] "Metron," Vol. 5 (1926), p. 90–104.

[2] See e.g. W. G. Madow, *Annals of Math. Stat.*, Vol. 11 (1940), p. 193.

In addition to the above mentioned test functions known as "variance within" and "variance among" classes other *symmetrical* test functions have been considered in the classical analysis of variance. Here again we may assume quite *general populations.* It results that the Lexis as well as the Poisson Series may now be characterized by *equalities* (instead of inequalities).

Finally it seems to be worthwhile to omit the assumption of independent chance variables and to study different kinds of *mutual* dependence. These investigations lead to new instructive *inequalities among the expectations.* These last considerations seem to be connected with Fisher's "intraclass correlation" and to supplement this idea.

On the Asymptotic Distribution of Medians of Samples from a Multivariate Population. A. M. Mood, University of Texas.

Let two variates x_1 and x_2 have a density function $f(x_1, x_2)$ which, besides being positive or zero and having its integral over the whole space equal to one, shall satisfy these conditions:

$$\int_{-\infty}^{\infty} f\left(x_1, \frac{1}{n}\right) dx_1 = \int_{-\infty}^{\infty} f(x_1, 0)\, dx_1 + O\left(\frac{1}{n}\right)$$

$$\int_{-\infty}^{\infty} f\left(\frac{1}{n}, x_2\right) dx_2 = \int_{-\infty}^{\infty} f(0, x_2)\, dx + O\left(\frac{1}{n}\right)$$

The coordinate system is assumed to have been chosen so that the population median is at the origin. Let $(\tilde{x}_1, \tilde{x}_2)$ be the median of a sample of $2n + 1$ elements drawn from a population with this density function. It is shown that for large samples $(\tilde{x}_1, \tilde{x}_2)$ is normally distributed to within terms of order $1/\sqrt{n}$ with zero means and variances and covariances given by certain integrals of $f(x_1, x_2)$.

A similar result is true for k as well as two variates.

A Problem in Estimation, Joseph F. Daly, The Catholic University of America.

Consider a normal population in which each individual is characterized by the variates $y_1, \cdots, y_p, y_{p+1}, y_{p+2}$. Suppose that the latter two are not directly observable, but that for given values of y_{p+1}, y_{p+2} the first set of y's is independently distributed about the "regression line" $y_k = y_{p+1} + ky_{p+2}$ $(k = 1, \cdots, p)$ with a common variance σ^2. For each individual, one can thus determine values $\hat{y}_{p+1}, \hat{y}_{p+2}$ from the observed $y_1, \cdots, y_p$, using the method of least squares. Assuming a similar relation between the expected values of $y_1, \cdots, y_{p+2}$ in the original population, these estimates $\hat{y}_{p+1}, \hat{y}_{p+2}$ are, of course, unbiased. However, if we calculate these $\hat{y}$'s for each individual of a sample of N, and substitute them in the Pearson product-moment correlation formula, the estimate of the correlation between y_{p+1} and y_{p+2} thus obtained is somewhat biased. The bias depends on the number of observable y's, and on the size of the variances and covariances of y_{p+1}, y_{p+2} relative to σ^2.

Is Sampling Theory Applicable to Economic Time Series? T. J. Koopmans, Penn Mutual Life Insurance Company.

The classical regression theory assumes that the values of the independent variables remain the same in repeated samples. Certain situations in economic analysis, like price formation according to the "cobweb" theorem, require a sampling theory of serial regression in which certain observations may represent a dependent variable at one time and an independent variable at a later time. This leads to the problem of the joint distribution of certain quadratic forms in normal variables.

The simplest problem of this type is that of the distribution of the ratio $r = q/p$ of a quadratic form q in T observations from a normal distribution with mean 0 to the sum p

of the squares of these observations. The distribution of r is independent of that of p and is

$$h(r) = \frac{\frac{1}{2}T - 1}{2\pi i} \int_{\gamma} \frac{(z - r)^{\frac{1}{2}T-2}}{\left\{\prod_{t=1}^{T} (z - k_t)\right\}^{\frac{1}{2}}} dz,$$

where the k_t are the characteristic values of q, while the path of integration γ proceeds from r through the lower half of the complex plane to a point on the real axis exceeding any k_t and from there returns to r through the upper half-plane.

In testing for the presence or absence of serial correlation (or regression) q is the sum of products of successive observations, and $k_t = \sigma^2 \cos \{\pi t/(T + 1)\}$. Replacing this set of discrete values in the above integral by a continuous variable of similar distribution, the following approximation to the distribution of r is found:

$$h^*(r) = \frac{T - 2}{\pi} 2^{\frac{1}{2}T} \int_{\text{arc sin } r}^{\frac{\pi}{2}} (\sin \phi - r)^{\frac{1}{2}T-2} \cdot \sin \left(\frac{T}{4} \pi - \frac{T + 1}{2} \phi\right) \cdot \cos^{\frac{1}{2}} \phi \, d\phi$$

CONSTITUTION

OF THE

INSTITUTE OF MATHEMATICAL STATISTICS

ARTICLE I

Name and Purpose

1. This organization shall be known as the Institute of Mathematical Statistics.
2. Its object shall be to promote the interests of mathematical statistics.

ARTICLE II

Membership

1. The membership of the Institute shall consist of Members, Fellows, Honorary Members, and Sustaining Members.

2. Voting members of the Institute shall be (a) the Fellows, and (b) all others who have been members for twenty-three months prior to the date of voting.

ARTICLE III

Officers, Board of Directors, Committee on Membership, and Committee on Publications

1. The Officers of the Institute shall be a President, two Vice-Presidents, and a Secretary-Treasurer, elected for a term of one year by a majority ballot at the annual meeting of the Institute. Voting may be in person or by mail.

(a) Exception. The first group of Officers shall be elected by a majority vote of the individuals present at the organization meeting, and shall serve until December 31, 1936.

2. The Board of Directors of the Institute shall consist of the Officers and the previous President.

3. The Institute shall have a Committee on Membership composed of three Fellows. At their first meeting subsequent to the adoption of this Constitution, the Board of Directors shall elect three members as Fellows to serve as the Committee on Membership, one member of the Committee for a term of one year, another for a term of two years, and another for a term of three years. Thereafter the Board of Directors shall elect from among the Fellows one member annually at their first meeting after their election for a term of three years. The president shall designate one of the Vice-Presidents as Chairman of this Committee.

4. The Institute shall have a Committee on Publications composed of three Members or Fellows elected by the Board of Directors. The President shall designate a Vice-President as Ex Officio Chairman of this Committee.

ARTICLE IV

Meetings

1. A meeting for the presentation and discussion of papers, for the election of Officers, and for the transaction of other business of the Institute shall be held annually at such time as the Board of Directors may designate. Additional meetings may be called from

time to time by the Board of Directors and shall be called at any time by the President upon written request from ten Fellows. Notice of the time and place of meeting shall be given to the membership by the Secretary-Treasurer at least thirty days prior to the date set for the meeting. All meetings except executive sessions shall be open to the public. Only papers accepted by a Program Committee appointed by the President may be presented to the Institute.

2. The Board of Directors shall hold a meeting immediately after their election and again immediately before the expiration of their term. Other meetings of the Board may be held from time to time at the call of the President or any two members of the Board. Notice of each meeting of the Board, other than the two regular meetings, together with a statement of the business to be brought before the meeting, must be given to the members of the Board by the Secretary-Treasurer at least five days prior to the date set therefor. Should other business be passed upon, any member of the Board shall have the right to reopen the question at the next meeting.

3. The Committee on Membership shall hold a meeting immediately after the annual meeting of the Institute. Further meetings of the Committee may be held from time to time at the call of the Chairman or any member of the Committee provided notice of such call and the purpose of the meeting is given to the members of the Committee by the Secretary-Treasurer at least five days before the date set therefor. Should other business be passed upon, any member of the Committee shall have the right to reopen the question at the next meeting.

4. At a regularly convened meeting of the Board of Directors, three members shall constitute a quorum. At a regularly convened meeting of the Committee on Membership, two members shall constitute a quorum.

ARTICLE V

Publications

1. The *Annals of Mathematical Statistics* shall be the Official Journal for the Institute. Other publications may be originated by the Board of Directors as occasion arises.

ARTICLE VI

Expulsion or Suspension

1. Except for non-payment of dues, no one shall be expelled or suspended except by action of the Board of Directors with not more than one negative vote.

ARTICLE VII

Amendments

1. This constitution may be amended by an affirmative two-thirds vote at any regularly convened meeting of the Institute provided notice of such proposed amendment shall have been sent to each voting member by the Secretary-Treasurer at least thirty days before the date of the meeting at which the proposal is to be acted upon. Voting may be in person or by mail.

BY-LAWS

ARTICLE I

Duties of the Officers, Board of Directors, Committee on Membership, and Committee on Publications

1. The President, or in his absence, one of the Vice-Presidents, or in the absence of the President and both Vice-Presidents, a Fellow selected by vote of the Fellows present,

shall preside at the meetings of the Institute and of the Board of Directors. At meetings of the Institute, the presiding officer shall vote only in the case of a tie, but at meetings of the Board of Directors he may vote in all cases. At least three months before the date of the annual meeting, the President shall appoint a Nominating Committee of three members. It shall be the duty of the Nominating Committee to make nominations for Officers to be elected at the annual meeting and the Secretary-Treasurer shall notify all voting members at least thirty days before the annual meeting. Additional nominations may be submitted in writing, if signed by at least ten Fellows of the Institute, up to the time of the meeting.

2. The Secretary-Treasurer shall keep a full and accurate record of the proceedings at the meetings of the Institute and of the Board of Directors, send out calls for said meetings and, with the approval of the President and the Board, carry on the correspondence of the Institute. Subject to the direction of the Board, he shall have charge of the archives and other tangible and intangible property of the Institute. He shall send out calls for annual dues and acknowledge receipt of same; pay all bills approved by the President for expenditures authorized by the Board or the Institute; keep a detailed account of all receipts and expenditures, prepare a financial statement at the end of each year and present an abstract of the same at the annual meeting of the Institute after it has been audited by a Member or Fellow of the Institute appointed by the President as Auditor. The Auditor shall report to the President.

3. The Board of Directors shall have charge of the funds and of the affairs of the Institute, with the exception of those affairs specifically assigned to the President or to the Committee on Membership. The Board shall have authority to fill all vacancies ad interim, occurring among the Officers, Board of Directors, or in any of the Committees. The Board may appoint such other committees as may be required from time to time to carry on the affairs of the Institute.

4. The Committee on Membership shall prepare and make available through the Secretary-Treasurer an announcement indicating the qualifications requisite for the different grades of membership.

5. The Committee on Publications, under the general supervision of the Board of Directors, shall have charge of all matters connected with the publications of the Institute, and of all books, pamphlets, manuscripts and other literary or scientific material collected by the Institute. Once a year this Committee shall cause to be printed in the Official Journal the Constitution and By-Laws and a classified list of all the Members and Fellows of the Institute.

ARTICLE II

Dues

1. Members shall pay five dollars at the time of admission to membership and shall receive the full current volume of the Official Journal. Thereafter, Members shall pay five dollars annual dues. The annual dues of Fellows shall be five dollars. The annual dues of Sustaining Members shall be fifty dollars. Honorary Members shall be exempt from all dues.

2. Annual dues shall be payable on the first day of January of each year.

3. The annual dues of a Fellow or Member include a subscription to the Official Journal. The annual dues of a Sustaining Member include two subscriptions to the Official Journal.

4. It shall be the duty of the Secretary-Treasurer to notify by mail anyone whose dues

may be six months in arrears, and to accompany such notice by a copy of this Article. If such person fail to pay such dues within three months from the date of mailing such notice, the Secretary-Treasurer shall report the delinquent one to the Board of Directors, by whom the person's name may be stricken from the rolls and all privileges of membership withdrawn. Such person may, however, be re-instated by the Board of Directors upon payment of the arrears of dues.

ARTICLE III

Salaries

1. The Institute shall not pay a salary to any Officer, Director, or member of any committee.

ARTICLE IV

Amendments

1. These By-Laws may be amended in the same manner as the Constitution or by a majority vote at any regularly convened meeting of the Institute, if the proposed amendment has been previously approved by the Board of Directors.

MEMBERS OF THE INSTITUTE OF MATHEMATICAL STATISTICS*

(*As of November 1, 1941*)

Acerboni, Dr. Argentino V. Banfield Larroque 232, Banfield, Argentina.

Alter, Prof. Dinsmore Director of Griffith Observatory, Los Angeles, Calif.

Anderson, Paul H. Ph.D. (Illinois) Dept. of Math., Louisiana State Univ., University, La.

Anderson, Richard L. Ph.D. (Iowa State Coll.) Part-time instr., North Carolina State Coll., Raleigh, N. C.

Anderson, Theodore W., Jr. B.S. (Northwestern) Instr., Princeton Univ., Princeton, N. J. *Graduate College.*

Anthony, Lucius Woodinville, Wash.

Arnold, Asso. Prof. H. E. Ph.D. (Yale) Wesleyan Univ., Middletown, Conn.

Arnold, Kenneth J. B.S. (Mass. Inst. of Tech.) 34 Field St., Boston, Mass.

Aroian, Leo A. Ph.D. (Michigan) Instr., Hunter Coll., New York, N. Y. *695 Park Ave.*

Arrow, Kenneth J. M.A. (Columbia) Fellow, Columbia Univ., New York, N. Y. *749 West End Ave.*

Ashcroft, A. Griffin M. E. (Cornell) Product Eng., Alex. Smith & Sons Carpet Co., Yonkers, N. Y.

Bachelor, Robert W. M.B.A. (Washington) American Bankers Association, New York, N. Y.

Bacon, Asst. Prof. Harold M. Ph.D. (Stanford) Stanford Univ., Stanford University, Calif. *Box 1144.*

Baker, George A. Ph.D. (Illinois) Experiment Sta., Coll. of Agric., Univ. of California, Davis, Calif..

Barnes, Jarvis M.A. (Peabody) Teacher, Atlanta City Schools, Atlanta, Ga. *744 Virginia Ct., NE.*

Barral-Souto, Prof. Jose Sc.D. (Univ. of Buenos Aires) Buenos Aires, Argentina. *Cordoba 1459.*

Barrett, Claudius S. M.A. (Northwestern) Dept. Chief, Western Electric Co., Inc., Hawthorne Sta., Chicago, Ill.

Bartky, Asso. Prof. Walter Ph.D. (Chicago) Univ. of Chicago, Chicago, Ill.

Baten, Asso. Prof. Walter D. Ph.D. (Michigan) Res. Asso., Mich. Agric. Exp. Sta., Mich. State Coll., East Lansing, Mich. *411 Marshall St..*

Bates, Prof. O. Kenneth Sc.D. (Mass. Inst. of Tech.) Cummings Prof. of Math. and Head of Dept., St. Lawrence Univ., Canton, N. Y.

Battin Isaac L. A.M. (Swarthmore) Instr., Brothers Coll. of Drew Univ., Madison, N.J.

Beal, Virginia B. B.A. (Mount Holyoke) Wis. Alum. Res. Found. Scholar, Univ. of Wisconsin, Madison, Wis. *230 North Brooks St.*

Beall, Dr. Geoffrey Dominion Entomological Lab., Chatham, Ont., Can. *729 Queen St.*

Bechhofer, Robert E. A.B. (Columbia) Jr. Statistician, Aberdeen Proving Ground, Aberdeen, Md. *107 Law St.*

---

* Members were asked to supply fresh information for this *Directory*. Records may be inexact or incomplete because of failure of some members to comply with this request. Changes in addresses, or errors in names, titles or addresses should be reported to the Secretary.

Been, Richard O. M.A. (George Washington) Agric. Economist, Bureau of Agric. Economics, Washington, D. C.
Bellinson, Harold R. M.S. (Mass. Inst. of Tech.) Asso. Statistician, Aberdeen Proving Ground, Aberdeen, Md. *Box 282.*
Bennett Prof. A. A. Ph.D. (Princeton) Brown Univ., Providence, R. I.
Benson Paul M.A. (Michigan) Instr., Bucknell Univ., Lewisburg, Pa.
Berger, Richard M.A. (Columbia) Nat. Bureau Econ. Research, New York, N. Y. *25 Rugby Rd., Rockville Centre, N. Y.*
Berkson, Dr. Joseph Mayo Clinic, Rochester, Minn.
Bernstein, Prof. Felix Ph.D. (Göttingen) New York Univ., New York, N. Y. *949 Broadway..*
Bingham, Marion D. A.B. (George Washington) Asst. Statistician, Dept. Agriculture, Washington, D. C. *325 N. Piedmont St., Arlington, Va.*
Blackburn, Asso. Prof. Raymond F. Ph.D. (Pittsburgh) Univ. of Pittsburgh, Pittsburgh, Pa.
Blake, Archie Ph.D. (Chicago) Asso. Mathematician, U. S. Coast and Geodetic Survey, Washington, D. C.
Blanche, Ernest E. Ph.D. (Illinois) Instr., Mich. State Coll., East Lansing, Mich.
Bliss, Chester I. Ph.D. (Columbia) Biometrician, New Haven and Storrs, Conn., Exp. Stations. *Conn. Agric. Exp. Station, Box 1106, New Haven, Conn.*
Boley, Charles C. M.S. (Illinois) Asst. Mining Eng., Ill. State Geol. Survey, Natural Resources Bldg., Urbana, Ill.
Bonis, Austin J. B.S. (C. C. N. Y.) 1225 Park Ave., New York, N. Y.
Book, Frances Bookkeeper, Salant and Salant, 56 Worth St., New York, N. Y. *2523 Aqueduct Ave., Bronx, N. Y.*
Boschan, Paul Ph.D. (Vienna) Statistician, Inst. of Applied Econometrics, 405 Lexington Ave., New York, N. Y.
Bowker, Albert H. S.B. (Mass. Inst. of Tech.) Res. Asst., Mass. Inst. of Tech. Cambridge, Mass.
Boyer, Prof. Lee E. Ed.D. (Penna. State Coll.) Millersville State Teachers Coll., Millersville, Pa.
Brady, Dorothy S. Ph.D. (California) Home Ec. Specialist, Bureau of Home Economics, Washington, D. C. *3848 Calvert St.*
Brandt, Alva E. Ph.D. (Iowa State Coll.) Prin. Soil Conservationist and Chief Conserv. Exp. Sta. Div., Washington, D. C. *Box 89, Route 3, Vienna, Va.*
Bridger C. A. M.S. (Oregon State Coll.) Statistician, State Dept. Public Health, Boise, Idaho. *P. O. Box 144, Eagle, Idaho.*
Bronfenbrenner, Martin Ph.D. (Chicago) Statistician and Analyst, Federal Reserve Bank of Chicago, Chicago, Ill.
Brookner, Ralph J. M.S. (Michigan) Ensign, Bureau of Ordnance, Navy Dept., Washington, D. C. *120 C St., NE, Apt. 202.*
Brooks, A. G. 2803 W. Erie St., Chicago, Ill.
Brown, George W. Ph.D. (Princeton) R. H. Macy and Co., New York, N. Y. *129 W. 86 St.*
Brown, Richard H. A.B. (Columbia) Foundation for Study of Cycles, Lexington Ave., New York, N. Y. *1107 John Jay Hall, Columbia Univ.*
Bryan, Joseph G. S.B. (Mass. Inst. of Tech.) 97 Green St., Melrose, Mass.
Burgess, R. W. Ph.D. (Cornell) Chief Statistician, Western Electric Co., 195 Broadway, New York, N. Y.
Burr, Asst. Prof. Irving W. Ph.D. (Michigan) Purdue Univ., W. Lafayette, Ind. *265 Littleten St.*
Bushey, Asso. Prof. J. H. Ph.D. (Michigan) Hunter Coll., New York, N. Y.
Caine, Walter E. M.B.A. (Northwestern) Sr. Rate Investigator, Fed. Power Com.; Dir.

Twentieth Century Fund, Survey of Rel. between Govt. and Elec. Light and Power Industry. *357 N. Glebe Rd., Arlington, Va.*

Calkins, Prof. Helen Ph.D. (Cornell) Head of Math. Dept., Penna. Coll. for Women, Pittsburgh, Pa.

Camp, Prof. Burton H. Ph.D. (Yale) Wesleyan Univ., Middletown, Conn. *110 Mt. Vernon St.*.

Carlson, John L. M.A. (Stanford) Instr., Reno High School, Reno, Nev. *736 West St.*

Carlton, A. George B.A. (Gustavus Adolphus) Jr. Statistician, War Dept. General Staff. G-4, Washington, D. C. *339 34th Pl., NE.*

Carver, Prof. H. C. Ph.D. (Michigan) Dept. of Math., Univ. of Michigan, Ann Arbor, Mich..

Cederberg, Prof. William E. Ph.D. (Wisconsin) Augustana Coll., Rock Island, Ill. *2542 22½ Ave.*

Cell, Asso. Prof. John W. Ph.D. (Illinois) North Carolina State Coll., Raleigh, N. C.

Chang, Z. T. A.M. (Columbia) 7 Lane 720, Avenue Foch, Shanghai, China.

Chalmers, Juanita 36 West 139 St., New York, N. Y.

Chapman, Roy A. U. S. Forest Service, 1051 New Federal Bldg., New Orleans, La.

Clark, Asso. Prof. A. G. A.M. (Colorado) Colorado State Coll. of A. and M., Fort Collins, Colo. *1012 Laporte Ave.*

Cochran, Prof. William G. M.A. (Cambridge) Iowa State Coll., Ames, Iowa..

Cohen, Capt. Alonzo C., Jr. Ph.D. (Michigan) Picatinny Arsenal, Dover, N. J.

Cohen, Melvin S. B.A. (Brooklyn) Jr. Statistician, Bureau of the Census, Washington, D. C. *100 C St., SE.*

Coleman, E. P. Dept. of Math., Municipal Univ. of Omaha, Omaha, Neb.

Coleman, Prof. James B. Ph.D. (California) Univ. of South Carolina, Columbia, S. C. *620 Bull St.*

Cotterman, Charles W. Ph.D. (Ohio State Univ.) Res. Asso., Univ. of Michigan, Ann Arbor, Mich.

Court, Louis M. M.A. (Cornell) 480 West 187 St., New York, N. Y.

Cowan, Donald R. G. Ph.D. (Minnesota) Mgr., Coml. Res. Dept., Republic Steel Corp., Republic Bldg., Cleveland, Ohio.

Cox, Gerald J. Ph.D. (Illinois) Sr. Commodity Specialist, Office of Prod. Mgt., Chem. Br., Washington, D. C. *3803 S St., NW.*

Cox, Prof. Gertrude M. M.S. (Iowa State Coll.) Head, Dept. of Experimental Statistics, North Carolina State Coll., Raleigh, N. C.

Craig, Asso. Prof. Allen T. Ph.D. (Iowa) Univ. of Iowa, Iowa City, Iowa..

Craig, Asso. Prof. Cecil C. Ph.D. (Michigan) Univ. of Michigan, Ann Arbor, Mich..

Crathorne, Prof. A. R. Ph.D. (Göttingen) Univ. of Illinois, Urbana, Ill..

Crowe, S. E. 137 University Dr., East Lansing, Mich.

Curtiss, Asst. Prof. John H. Ph.D. (Harvard) Cornell Univ., Ithaca, N. Y.

Daly, Joseph F. Ph.D. (Princeton) Instr., Catholic Univ. of Am., Washington, D. C.

Dantzig, George B. 2609 Fulton St., Berkeley, Calif.

Davies, Prof. George R. Ph.D. (North Dakota) Univ. of Iowa, Iowa City, Iowa.

Day, Besse B. A.M. (Michigan) Statistician, Forest Serv., Dept. of Agriculture, Washington, D. C.

Deemer, Walter L., Jr. A.B. (Lehigh) Educ. Res. Corp., 13 Kirkland St., Cambridge, Mass.

De Lury, Daniel B. Ph.D. (Toronto) Lecturer, Univ. of Toronto, Toronto, Ont., Can.

Deming, W. Edwards Ph.D. (Yale) Prin. Mathematician, Bur. of the Census, Washington, D. C..

Di Salvatore, Philip M.A. (Princeton) Head, Reinsurance Sec., Guardian Life Ins. Co., New York, N. Y.

Dixon, Wilfrid J. M.A. (Wisconsin) Res. Asst., Princeton Univ., Princeton, N. J. *35 Humbert St.*

Dodd, Prof. Edward L. Ph.D. (Yale) Univ. of Texas, Austin, Texas *3012 West Ave.*.

Dodge, Harold F. A.M. (Columbia) Quality Results Eng., Bell Telephone Labs., Inc., New York, N. Y.

Doob, Asst. Prof. J. L. Ph.D. (Harvard) Univ. of Illinois, Urbana, Ill. *43 Linden Lane, Princeton, N. J.*.

Dorn, Harold F. Ph.D. (Wisconsin) Sr. Statistician, U. S. Public Health Serv., Washington, D. C. *15 Burning Tree Ct., Bethesda, Md.*

Dorweiler, Paul B.S. (Iowa) Actuary, Aetna Casualty and Surety Co., Hartford, Conn.

Dresch, Lt. (jg) Francis W. Ph.D. (California) U. S. N. R., Naval Proving Gr., Dahlgren, Va.

Dressel, Asst. Prof. Paul L. Ph.D. (Michigan) Michigan State Coll., East Lansing, Mich.

Dunlap, Jack W. Ph.D. (Columbia) Catherine Strong Hall, Univ. of Rochester, Rochester, N. Y.

Durand, David A.M. (Columbia) Natl. Bur. of Economic Res., Hillside, Riverdale, N. Y.

Dutka, Jacques A.M. (Columbia) 56 West 180 St., Bronx, N. Y.

Dwyer, Asst. Prof. Paul S. Ph.D. (Michigan) Univ. of Michigan, Ann Arbor, Mich. *2509 James St.*.

Edgett, G. L. Queen's Univ., Kingston, Ont., Can.

Eisenhart, Asst. Prof. Churchill Ph.D. (London) Univ. of Wisconsin, Madison, Wis. *Tower Room, Agronomy Bldg.*

Elkin, William F. M.S. (Michigan) Kellogg Found. Fellow in Pub. Health Statistics, Allegan County Health Dept., Allegan, Mich.

Elkins, Thomas A. A.M. (Princeton) Geophysicist, Gulf Res. and Dev. Co., Pittsburgh, Pa. *3255 Parkview Ave.*

Elston, James S. Travelers Insurance Co., Hartford, Conn.

Elting, John P. Kendall Mills, Paw-Creek, N. C.

Elveback, Mary L. M.A. (Minnesota) Hunter Coll., Park Ave. and 68 St., New York, N. Y.

Embody, Daniel R. M.S. (Cornell) Instr., Cornell Univ., Ithaca, N. Y.

Ettinger, Wallace J. B.S. (Lewis Inst.) Designing Eng., Edison Gen. Elec. Appliance Co., 5600 West Taylor St., Chicago, Ill.

Eudey, Mark A.B. (California) 1904 University Ave., Apt. 4, Berkeley, Calif.

Evans, Asso. Prof. Herbert P. Ph.D. (Wisconsin) Univ. of Wisconsin, Madison, Wis.

Evans, W. D. B.S. (Clarkson) Prin. Economist, Bur. of Labor Statistics, Dept. of Labor, Washington, D. C.

Faust, Richard H. B.S. (Yale) U. S. Army, Co. A. 28th Inf. Trn. Btn., Camp Croft, S. C. *239 Vreeland Ave., Nutley, N. J.*

Feldman, Hyman M. Ph.D. (Washington) Teacher, Beaumont H. S., St. Louis, Mo.

Feller, Asst. Prof. Willy Ph.D. (Göttingen) Brown Univ., Providence, R. I.

Fertig, Prof. John W. Ph.D. (Minnesota) Columbia Univ., De Lamar Inst. of Pub. Health, 600 West 168 St., New York, N. Y.

Fischer, Carl H. Ph.D. (Iowa) Math. Dept., Univ. of Mich., Ann Arbor, Mich.

Fisher, Prof. Irving Yale University, New Haven, Conn. *460 Prospect St.*.

Flaherty, Asst. Prof. William C. A.B. (Georgetown) Georgetown Univ., Washington, D. C.

Flood, M. M. Ph.D. (Princeton) Res. Asso., Princeton Local Govt. Survey, Princeton, N. J.

Foster, Ronald M. S.B. (Harvard) Bell Telephone Labs., 463 West St., New York, N. Y. *122 East Dudley Ave., Westfield, N. J.*

Fox, Asso. Prof. Philip G. A.M. (Wisconsin) Univ. of Wisconsin, Madison, Wis.

Frankel, Lester R. 1445 Otis Pl., NW, Washington, D. C.

Freeman, Asst. Prof. Harold A. S.B. (Mass. Inst. of Tech.) Mass. Inst. of Tech., Cambridge, Mass.

Friedman, Milton M.A. (Chicago) Consult. Expert, Treasury Dept., Washington, D. C.; Natl. Bur. of Econ. Res., 1819 Broadway, New York, N. Y. *2480 16 St., NW, Washington, D. C.*

Fry, Thornton C. Ph.D. (Wisconsin) Res. Mathematician, Bell Telephone Labs., 463 West St., New York, N. Y..

Fryer, Asst. Prof. Holly C. Ph.D. (Iowa State Coll.) Kansas State Coll., Manhattan, Kan.

Gage, Robert P. M.S. (Iowa State Coll.) Asso. in Dept. of Biometry and Med. Statistics, Mayo Clinic, Rochester, Minn.

Gause, G. Rupert B.S. (Citadel) Asst. Statistician, War Dept., Aberdeen Proving Gr., Md. *Box 179, Aberdeen, Md.*

Geiringer, Hilda P. Ph.D. (Vienna) Lecturer, Bryn Mawr Coll., Bryn Mawr, Pa..

Gibson, Robert W. A.M. (Illinois) 602 S. Busey St., Urbana, Ill.

Gintzler, Leone B. M.A. (California) Stat. Res. Asst., Bur. of Pub. Adm., Univ. of California, Berkeley, Calif.

Girshick, M. A. Bur. of Home Economics, Dept. of Agriculture, Washington, D. C.

***Glover, Prof. James W.** Dept. of Math., Univ. of Michigan, Ann Arbor, Mich..

Gordon, Robert D. A.M. (Stanford) Res. Asst., Scripps Inst. of Oceanography, La Jolla, Calif.

Graves, Asst. Prof. Clyde H. Ph.D. (Chicago) Penna. State Coll., State College, Pa.

Greenwood, Asst. Prof. Joseph A. Ph.D. (Missouri) Duke Univ., Durham, N. C. *1314 Norton St.*

Greville, Thomas N. E. Ph.D. (Michigan) Asso. Actuarial Mathematician, Bur. of the Census, Washington, D. C.

Griffin, John I. Ph.D. (Columbia) Instr., Long Island Univ., Brooklyn, N. Y. *115 Henry St.*

Grove, Asst. Prof. C. C. Ph.D. (Johns Hopkins) Coll. of the City of New York, 17 Lexington Ave., New York, N. Y. *143 Milburn Ave., Baldwin, N. Y.*

Guard, Harris T. M.S. (Colorado) Instr., Colorado State Coll., Fort Collins, Colo.

Gumbel, Asso. Prof. Emil J. Ph.D. (Munich) New School for Social Res., 66 West 12 St., New York, N. Y. *3820 Waldo Ave..*

Guttman, Louis M.A. (Minnesota) Instr., Cornell Univ., Ithaca, N. Y.

Hagood, Mrs. Margaret Jarman Ph.D. (North Carolina) Res. Asso., Univ. of North Carolina, Chapel Hill, N. C. *One Village Apts.*

Haines, Harold M.S. (New York) Res. Asst., Burndy Engineering Co., 459 East 133 St., New York, N. Y.

Hammer, Preston C. Ph.D. (Ohio State Coll.) Instr., Oregon State Coll., Corvallis, Ore.

Hand, Howard J. B.S. (Carnegie Inst. of Tech.) Head, Stat. Div., Met. Dept., Natl. Tube Co., Lorain, Ohio *628 Lakeside Ave.*

Hansen, Morris H. M. A. (American) Acting Asst. Chief Stat., Bur. of the Census, Washington, D. C.

Harshbarger, Asst. Prof. Boyd M.A. (Illinois) Virginia Polytechnic Inst., Blacksburg, Va. *Fellows Reading Room, George Washington Univ., Washington, D. C.*

Hart, Bertha I. A.M. (Cornell) Sr. Computer, Ballistic Res. Lab., Aberdeen Proving Gr., Md.

Hart, Prof. William L. Ph.D. (Chicago) Univ. of Minnesota, Minneapolis, Minn.

Head, George A. Technical Staff, Bell Telephone Labs., Inc., 463 West St., New York, N. Y.

Hebley, Henry F. Product Control Mgr., Pittsburgh Coal Co., P.O. Box 145, Pittsburgh, Pa.

---

* Deceased.

Heidingsfield Myron S. Ph.D. (New York) Instr., New York Univ., New York, N. Y.; Director, Survey of Res. in Recreation, Fed. Works Agency, New York, N. Y. *172 West 79 St.*

Henderson, Robert D.Sc. (Toronto) Vice Pres. and Actuary, Retired, Equitable Life Assur. Soc., New York, N. Y. *Crown Point, Essex Co., N. Y.*.

Hendricks, Walter A. Div. of Agr. Stat., Agr. Marketing Serv., U. S. Dept. of Agriculture, Washington, D. C.

Henry, Malcolm H. M.S. (Michigan) Asst. Stat., Michigan Dept. of Soc. Welfare, Lansing, Mich. *1726 Todd Ave.*

Hermie, Albert M. M.S. (Illinois) Jr. Statistician, Office of Prod. Mgt., 1102 Raleigh Hotel, Washington, D. C. *Fellows Reading Room, George Washington Univ.*

Hildebrandt, Asst. Prof. E. H. C. Ph.D. (Michigan) State Teachers Coll., Upper Montclair, N. J.

Hinds, Asst. Prof. Frances Campbell (Mrs.) M.A. (California) George Peperdine Coll., Los Angeles, Calif. *1121 West 79 St.*

Hoel, Asst. Prof. Paul G. Ph.D. (Minnesota) Univ. of California, Los Angeles, Calif.

Hopper, Asst. Prof. Grace M. Ph.D. (Yale) Vassar Coll., Poughkeepsie, N. Y.

Horst, Aaron P. Ph.D. (Chicago) Supervisor of Selection Res., Procter & Gamble, Cincinnati, Ohio.

Hotelling, Prof. Harold Ph.D. (Princeton) Columbia Univ., New York, N. Y. *Mountain Lakes, N. J.*.

Householder, Alston S. Ph.D. (Chicago) Res. Asso. in Math. Biophysics, Univ. of Chicago, Chicago, Ill.

Hoy, Asst. Prof. Elvin A. Univ. of Hawaii, Honolulu, Hawaii *509 West 121 St., New York, N. Y.*

Hsu, Chung-Tsi M.S. (London) 206 Livingston Hall, Columbia Univ., New York, N. Y.

Huntington, Prof. Emeritus Edward V. Ph.D. (Strassburg), S. D. (San Marcos) Harvard Univ., Cambridge, Mass. *48 Highland St.*.

Hurwitz, William 119 Concord Way, Washington, D. C.

Ingraham, Prof. Mark H. Ph.D. (Chicago) Univ. of Wisconsin, Madison, Wis..

Jablon, Seymour A.M. (Columbia) 200 West 108 St., New York, N. Y.

Jackson, Prof. Dunham Ph.D. (Göttingen) Univ. of Minnesota, Minneapolis, Minn..

Jackson, Robert W. B. Ph.D. (London) Lecturer, Univ. of Toronto, Toronto, Ont., Can.

Jacobs, Walter A.M. (George Washington) Asst. Finan. Stat., Securities and Exchange Commission, Washington, D. C. *1421 Somerset Pl.*

Juran, J. M. Western Electric Co., 195 Broadway, New York, N. Y.

Jaramillo, Trinidad J. Ph.D. (Chicago) Actuary, Bur. of the Treasury; Instr., Far Eastern Univ., Manila, P. I. *P. O. Box 1045.*

Johner, Paul B.S. (Carnegie Inst. of Tech.) First Lt., Ordnance Dept., Pittsburgh Ordnance Dist., 1202 Chamber of Commerce Bldg., Pittsburgh, Pa. *1118 Victoria Ave., New Kensington, Pa.*

Johnson, Asst. Prof. Evan, Jr. Ph.D. (Chicago) Penna. State Coll., State College, Pa.

Johnson, Prof. Palmer O. Ph.D. (Minnesota) Univ. of Minnesota, Minneapolis, Minn.

Kantorovitz, Myron Ph.D. (Berlin) Res. Fellow, Milbank Memorial Fund, 40 Wall St., New York, N. Y. *2446 29 Street, Astoria, N. Y.*

Katz, Leo M. A. (Wayne) Statistician, Dept. of Labor and Industry, Lansing, Mich.

Katz, Mortimer B. M.A. (George Washington) Jr. Statistician, Bur. of the Census, Washington, D. C.

Katzoff, E. Taylor Ph.D. (Northwestern) Asst., Northwestern Univ., Evanston, Ill.

Keeping, E. S. Dept. of Math., Univ. of Alberta, Edmonton, Alberta, Can.

Keffer, Ralph M.A. (Wisconsin) Aetna Life Ins. Co., Hartford, Conn.

Kelley, Prof. Truman L. Ph.D. (Columbia) Grad. Sch. of Educ., Peabody House, Kirkland St., Cambridge, Mass..

Kenney, John F. A.M. (Michigan) Univ. Extension, Univ. of Wisconsin, Milwaukee, Wis.

Kent, Robert H. A.M. (Harvard) Asso. Director, Ballistic Res. Lab., Aberdeen Proving Gd., Md.

Kiernan, Prof. Charles J. M.S. (Columbia) St. John's Univ., Brooklyn, N. Y. *14 Fairbanks St., Hillside, N. J.*

Kimball, Bradford F. Ph.D. (Cornell) Sr. Statistician, N. Y. State Pub. Serv. Commission, 80 Centre St., New York, N. Y. *33 Bogart Ave., Port Washington, N. Y.*

King, Arnold J. B.S. (Wyoming) Agr. Statistician, U. S. D. A., A.M.S., Statistical Lab., Iowa State Coll., Ames, Iowa.

King, Frederick G. A.B. (Harvard) Asst. Statistician, N. Y. State Pub. Serv. Commission, New York, N. Y. *609 West 115 St.*

Kingston, Prof. Jorge C. E. (Brazil) Univ. of Brazil, Rio de Janeiro, Brazil *44 Rua Rita Ludolf.*

Knowler, Asst. Prof. L. A. Ph.D. (Iowa) Univ. of Iowa, Iowa City, Iowa.

Knudsen, Lila F. B.S. (Minnesota) Asst. Mathematician, Food and Drug Adm., Fed. Sec. Agency, Washington, D. C.

Kohl, Alma 1209 W. Sherwin Ave., Chicago, Ill.

Konijn, Hendrik S. C.E.W. (Rotterdam) Statistician, Natl. Bur. of Econ. Res., West 254 St. and Independence Ave., New York, N. Y. *245 East 72 St.*

Koopmans, Tjalling Dr. Math. & Phy. Sc. (Leiden) Penn Mutual Life Ins. Co., Philadelphia, Pa.

Kossack, Carl F. Ph.D. (Michigan) Instr., Univ. of Oregon, Eugene, Ore.

Kozelka, Asst. Prof. Richard L. Ph.D. (Minnesota) Univ. of Minnesota, Minneapolis, Minn.

Kullback, Solomon Ph.D. (George Washington) Lecturer, George Washington Univ., Washington, D. C..

Kurtz, Albert K. Ph.D. (Ohio State) Res. Asso., Life Ins. Sales. Res. Bur., Hartford, Conn. *173 Cornwall St.*

Kury, Anita R. M.A. (Michigan) Analyst, Adm. Staff, W. P. A., 70 Columbus Ave., New York, N. Y. *188 Beach 63 St., Arverne, N. Y.*

Kwerel, Seymour M. B.S. (C. C. N. Y.) Econ. Analyst and Stat., Bur. of For. & Dom. Commerce, Dept. of Commerce, Washington, D. C. *1556 Minford Pl., Bronx, N. Y.*

Laderman, Jack 3224 Bronx Blvd., Bronx, N. Y.

Lancaster, Asst. Prof. Otis E. Ph.D. (Harvard) Univ. of Maryland, College Park, Md.

Lange, Asso. Prof. Oscar Ll.D. (Cracow) Univ. of Chicago, Chicago, Ill.

Larsen, Asst. Prof. Harold D. Ph.D. (Wisconsin) Univ. of New Mexico, Albuquerque, N. M.

Leavens, Dickson H. M.A. (Yale) Res. Asso., Cowles Comm. for Res. in Economics, Univ. of Chicago, Chicago, Ill.

Le Leiko, Max B.S. (New York) Statistician, Williard Contracting Co., 221 West 57 St., New York, N. Y.

Lemme, Asst. Prof. Maurice M. A.M. (Indiana) Univ. of Toledo, Toledo, Ohio

Lengyel, Bela A. Ph.D. (Pázmány) Instr., Rensselaer Poly. Inst., Troy, N. Y.

Levin, Ida M.Sc. (Johns Hopkins) 118 Millard Hall, Univ. of Minnesota, Minneapolis, Minn.

Levin, Pvt. Joseph H. Ph.D. (Chicago) U. S. Army, Coast Art. Sch. Detachment, Fort Monroe, Va., *12515 Broadstreet Blvd., Detroit, Mich.*

Livers, Asst. Prof. Joe J. M.A. (Washington State Coll.) Montana State Coll., Bozeman, Mont.

Livesay, Naomi Ph.M. (Wisconsin) Rock. Found. Fellow, Univ. of Chicago, Chicago, Ill. *6010 Dorchester Ave.*

Lorge, Asso. Prof. Irving Ph.D. (Columbia) Columbia Univ., New York, N. Y.

Lotka, Dr. A. J. Metropolitan Life Ins. Co., One Madison Ave., New York, N. Y..

Lukacs, Eugene Ph.D. (Vienna) 5510 Pimlico Rd., Baltimore, Md.
Lundberg, Prof. George A. Ph.D. (Minnesota) Bennington Coll., Bennington, Vt.
McCarthy, Michael D. Univ. Coll., Cork, Ire.
McCarthy, Philip J. A.M. (Princeton) Asst., Princeton Univ., Princeton, N. J.
McDiarmid, Asst. Prof. Orville J. Ph.D. (Harvard) Carnegie Inst. of Tech., Pittsburgh, Pa.
McEwen, Prof. G. F. Ph.D. (Stanford) Scripps Institution, La Jolla, Calif.
Macphail, Asst. Prof. Moray St. J. Ph.D. (Oxford) Acadia Univ., Wolfville, N. S. *4 Queenston Pl., Princeton, N. J.*
Madow, William G. Ph.D. (Columbia) Bur. of the Census, Washington, D. C. *1445 Ogden St., NW, Washington, D. C.*.
Malzberg, Benjamin Ph.D. (Columbia) Sr. Statistician, N. Y. Dept. of Mental Hygiene, Albany, N. Y.
Mansfield, Prof. Ralph S.M. (Chicago) Chicago Teachers Coll., 6800 S. Stewart Ave., Chicago, Ill.
Marschak, Prof. Jakob Ph.D. (Heidelberg) New School for Social Res., 66 West 12 St., New York, N. Y.
Marcuse, Mrs. Sophie M.A. (Columbia) *218 18 St., Santa Monica, Calif.*
Mauchly, John W. Ph.D. (Johns Hopkins) Instr., Univ. of Pennsylvania, Philadelphia, Pa.
Merrell, Margaret Sc.D. (Johns Hopkins) Asso., Sch. of Hyg. & Pub. Health, Johns Hopkins Univ., Baltimore, Md.
Michie, Prof. James N. M.A. (Michigan) Head, Math. Dept., Texas Tech. Coll., Lubbock, Tex. *Box 91, Texas Tech. Sta.*
Mills, Vicente Box 2090, Manila, P. I.
Milroy, Prof. Robert R. Ph.D. (Iowa) Bowling Green Coll. of Commerce, Bowling Green, Ky.
Miner, John R. Sc.D. (Johns Hopkins) Asso. Editor, Mayo Clinic, Rochester, Minn. *619 Seventh Ave., SW.*
Mises, Prof. Richard v. Dr. (Vienna) Harvard Univ., Cambridge, Mass..
Mode, Prof. Elmer B. A.M. (Harvard) Boston Univ., Boston, Mass.
Molina, Edward C. Bell Telephone Labs., 463 West St., New York, N. Y..
Mood, Alexander M. Ph.D. (Princeton) Instr., Univ. of Texas, Austin, Tex.
Moore, Marjorie E. B.A. (Macalester) Asst., Univ. of Minnesota, Minneapolis, Minn. *727 S. E. 15 Ave.*
Morrison, Nathan A.B. (Brooklyn) Sr. Statistician, N. Y. Div. of Place. & Unemploy. Ins., 342 Madison Ave., New York, N. Y. *420 Avenue F, Brooklyn, N. Y.*
Morton, Asso. Prof. Joseph E. D.Sc. (Geneva), J. D. (Charles) Knox Coll., Galesburg, Ill.
Mosteller, Frederick C. M.S. (Carnegie Inst. of Tech.) Asst., Princeton Univ., Princeton, N. J. *26 Bank St.*
Mouzon, Prof. Edwin D., Jr. Ph.D. (Illinois) Head, Math. Dept., Southern Methodist Univ., Dallas, Tex.
Mudgett, Prof. Bruce D. Ph.D. (Pennsylvania) Univ. of Minnesota, Minneapolis, Minn.
Muhlendorf, Mildred 3424 Gates Pl., Bronx, N. Y.
Mummery, Charles R. B.A. (Toronto) Sec. Head, The Hoover Co., N. Canton, Ohio *606 E. Maple St.*
Munch, Adabelle B. 8070 87 Ave., Woodhaven, N. Y.
Naugle, Leah A.M. (Michigan) Statistical Clerk, Grad. School, Univ. of Michigan, Ann Arbor, Mich. *414 S. Division St.*
Neifeld, Morris R. Ph.D. (New York) Economist, 15 Washington St., Newark, N. J.
Nekrassoff, Vladimir A. Dr. Eng. (Micharlovskaya Act. Acad.) Asst. Statistician, Ballistic Res. Lab., Aberdeen Proving Ground, Md. *Box 5, Havre de Grace, Md.*
Nesbitt, Dr. C. J. Univ. Michigan, Ann Arbor, Mich.

Netzer, Bennie M.S. (Carnegie Inst. of Tech.) 2521 Sarah St., Pittsburgh, Pa.
Neyman, Prof. Jerzy Ph.D. (Warsaw) Dir. Statistical Laboratory, Univ. of California, Berkeley, Calif..
Nisselson, Harold B.S. (C. C. N. Y.) Asst. Statistician, Bur. of the Census, Washington, D. C. *1737 Q St., NW.*
Niver, George E. Mgr., Coml. Res. Dept., Natl. Tube Co., Pittsburgh, Pa.
Norden, Monroe L. S.B. (Mass. Inst. of Tech.) Jr. Statistician, Aberdeen Proving Ground, Md. *Aberdeen, Md.*
Norris, Nilan Ph.D. (Stanford) Instr., Hunter Coll., 695 Park Ave., New York, N. Y. *2800 University Ave.*
Northam, Jack I. M.A. (Mich. State Coll.) Teach. Fel., Univ. Michigan, Ann Arbor, Mich.
Norton, Kenneth A. B.S. (Chicago) Sr. Radio Eng., Fed. Commun. Comm., Washington, D. C. *4623 Kenmore Dr., NW.*
O'Callahan, Rev. Joseph T. S.T.L. (Gregorian) Prof. of Math., Holy Cross Coll., Worcester, Mass.
Olds, Asso. Prof. Edwin G. Ph.D. (Pittsburgh) Carnegie Inst. of Tech., Pittsburgh, Pa. *953 La Clair Ave., Regent Square..*
Ollivier, Dr. Arthur Box 405, State College, Miss.
Olmstead, Paul S. Ph.D. (Princeton) Tech. Staff, Bell Telephone Labs., 463 West St., New York, N. Y. *Box 72, Essex Fells, N. J.*
Olshen, Abraham C. Ph.D. (Iowa) Actuary & Chief Examiner, Oregon Ins. Dept., Salem, Ore.
Ore, Prof. Oystein Ph.D. (Oslo) Yale Univ., New Haven, Conn.
Osborne, James G. B.S. (California) Chief, Sec. Forest Meas., Forest Service, Washington, D. C.
O'Toole, Alphonsus L. Ph.D. (Michigan) Fed. Commun. Comm., Washington, D. C..
Owen, Ruth L. M.A. (Columbia) Rayon Production Accountant, E. I. Du Pont de Nemours & Co., Waynesboro, Va. *201 N. Coalter St., Staunton, Va.*
Ozanne, Paul Univ. of Wisconsin, Madison, Wis.
Passano, Russell F. B.S. (Johns Hopkins) Res. Eng., Bethlehem Steel Co., Bethlehem, Pa. *209 Wall St.*
Paulsen, Edward M.A. (Columbia) Jr. Statistician, New York Bd. of Ed., 110 Livingston St., Brooklyn, N. Y. *432 W. 34 St., New York, N. Y.*
Payne, Asst. Prof. Charles K. Ph.D. (New York) Washington Sq. Coll., New York Univ., New York, N. Y.
Peach, Paul Asst. Ed., "The Freeman," Freeman Corp., 30 E. 29 St., New York, N. Y.
Peiser, Pvt. Donald E. M.A. (Columbia) U. S. Army, Camp Upton, Long Island, N. Y. *1117 New York Ave., Brooklyn, N. Y.*
Perlo, Victor A.M. (Columbia) 5426 Commerce Bldg., Washington, D. C.
Perotti, Juliette M. B.A. (Hunter) 5021 205 St., Bayside, N. Y.
Perry, Prof. Robert D. Ph.D. (Peabody) Statistician, Baylor Univ., Waco, Tex.
Perryman, James H. 12 Richmond Ave., Ilfracombe, Devon, Eng.
Peterson, Asst. Prof. Andrew I. M.S. (Columbia) New York Univ., Univ. Hts., New York, N. Y.
Petit, Lucette B.A. (Hunter) 4216 80 St., Elmhurst, N. Y.
Petrie, George W., III M.S. (Carnegie Inst. of Tech.) Pittsburgh Piping & Equipment Co., Pittsburgh, Pa. *613 Washington Ave., Oakmont, Pa.*
Pierce, Prof. Joseph A. Ph.D. (Michigan) Atlanta Univ., Atlanta, Ga.
Pines, Sylvia F. M.A. (Michigan) Jr. Statistician, Natl. Def. Coun., Washington, D. C. *1235 Randolph St., NW.*
Pixley, Asso. Prof. Henry H. Ph.D. (Chicago) Wayne Univ., Detroit, Mich.
Pollard, Asso. Prof. H. S. Ph.D. (Wisconsin) Miami Univ., Oxford, O.

Preinreich, Gabriel A. D. Ph.D. (Columbia) Cert. Pub. Accountant, 17 E. 42 St., New York N. Y.

Press, Harry B.A. (Brooklyn) 1801 55 St., Brooklyn, N. Y.

Ratkowitz, Elsie 3280 Spencer Dr., Bronx, N. Y.

Reed, Prof. Lowell J. Ph.D. (Pennsylvania) Johns Hopkins Univ., Baltimore, Md. *615 N. Wolfe St.*

Regan, Asso. Prof. Francis Ph.D. (Michigan) St. Louis Univ., St. Louis, Mo.

Reitz, Asst. Prof. Wilhelm Ph.D. (Wisconsin) Examiner, Coll. Ed., Wayne Univ., Detroit, Mich.

Reno, Franklin V. A.M. (Virginia) Box 423, Aberdeen, Md.

Richardson, Prof. C. H. Ph.D. (Michigan) Bucknell Univ., Lewisburg, Pa.

Rider, Prof. Paul R. Ph.D. (Yale) Washington Univ., St. Louis, Mo..

Rietz, Prof. Henry L. Ph.D. (Cornell) Head, Math. Dept., Univ. of Iowa, Iowa City, Iowa..

Risley, Adj. Prof. A. Marguerite A.M. (Syracuse) Randolph-Macon Woman's Coll., Lynchburg, Va.

Roberts, Clifford Dir. of Research, United Shoe Machinery Co., 140 Federal St., Boston, Mass.

Robinson, Selby L. Ph.D. (Iowa) Coll. City of New York, New York, N. Y. *4138 W. Moreland St., Little Neck, N. Y.*

Romig, H. G. Bell Telephone Labs., 463 West St., New York, N. Y.

Roodkowsky, Nikita D. A.M. (Columbia) Consul. Stat., H. A. Hopf Co., 500 Fifth Ave., New York, N. Y. *310 W. 109 St.*

Roos, C. F. Ph.D. (Rice) Dir. of Research, Inst. of Appl. Econometrics, 420 Lexington Ave., New York, N. Y..

Rosander, A. C. Ph.D. (Chicago) Statistician, Office of Prod. Mgt., Social Security Bldg., Washington, D. C. *2431 E St., NW.*

Ross, Norman W. Fel. Inst. Act. Refuge Assurance Co., Oxford St., Manchester, Eng.

Royston, Asst. Prof. Robert W. M.S. (Michigan) Washington & Lee Univ., Lexington, Va.

Rubin, Ernest M.A. (Columbia) 3073 Williamsbridge Rd., New York, N. Y.

Ruger, Asso. Prof. Henry A. Ph.D. (Columbia) Retired, Columbia Univ., New York, N. Y. *526 N. Main St., Wellington, Ohio.*

Rulon, Asst. Prof. Philip J. Ph.D. (Minnesota) Harvard Univ., Cambridge, Mass. *13 Kirkland St.*

Sachs, Rose B.A. (Goucher) 2800 Reisterstown Rd., Baltimore, Md.

Sandomire, Marion M. A.B. (Hunter) Stat. Ed., Office of Information, Dept. of Agriculture, Washington, D. C.

Sasuly, Max M.S. (Chicago) Social Security Bd., Washington, D. C. *3815 Upton St.*

Satterthwaite, Franklin E. Ph.D. (Iowa) Group Stat. Div., Aetna Life Ins. Co., Hartford, Conn.

Savulak, Asso. Prof. John H. M.A. (Duquesne) Duquesne Univ., Pittsburgh, Pa. *1334 Oakhill St., NS.*

Scheffé, Henry Ph.D. (Wisconsin) Instr., Princeton Univ., Princeton, N. J.

Schrock, Edward M. B.S. (Pittsburgh) Sup., Stat. Div., Metal. Dept., Jones & Laughlin Steel Corp., Aliquippa, Pa.

Schug, Howard L. B.S. (Lafayette) Res. Eng., 806 Portage St., North Canton, O.

Schultz, 1st Lt. Andrew S., Jr. Ph.D. (Cornell) U. S. Army, Of., Chief of Ord., Washington, D. C.; Asst. Prof. (on leave), Cornell Univ., Ithaca, N. Y. *Apt. 230, Arlington Village, Arlington, Va.*

Schwartz, Herman M. Ph.D. (Pennsylvania) Instr., Univ. of Illinois, Urbana, Ill.

Scott, Elizabeth L. A.B. (California) Res. Asst., Univ. of California, Berkeley, Calif. *34 Tunnel Rd.*

Seeley, Burton D. Bur. of Labor Statistics, Washington, D. C.

Seiden, Nathan M.P.A. (New York) Statistician, New York City Housing Authority, 122 E. 42 St., New York, N. Y. *1520 Shakespeare Ave., Bronx, N. Y.*

Sells, Saul B. Ph.D. (Columbia) Program Analyst, Fed. Works Ag., Washington, D. C.; Instr., Brooklyn Coll., Brooklyn, N. Y. *3555 168 St., Flushing, N. Y.*

Shaw, Lawrence W. A.M. (Pennsylvania) Asst. Statistician, Ballistic Res. Lab., Aberdeen Proving Gd., Md.

Shelton, William C. A.B. (Princeton) Asso. Statistician, Dept. of Commerce, Washington, D. C. *2436 N. Columbus St., Arlington, Va.*

Shen, Eugene Ph.D. (Stanford) 730/556 Connaught Rd., Shanghai, China

Shephard, Ronald W. Ph.D. (California) Instr.-Res. Fel., Univ. of California, Berkeley, Calif. *2511 Hearst Ave.*

Shewhart, Walter A. Ph.D. (California) Eng., Bell Telephone Labs., 463 West St., New York N. Y. *158 Lake Dr., Mountain Lakes, N. J.*.

Shohat, Asso. Prof. J. A. Ph.D. (Petrograd) Math. Dept., Univ. of Pennsylvania, Philadelphia, Pa..

Siller, Harry M.S. (New York) Jr. Stat., 42 Independence Ave., SW, Washington, D. C.

Simmons, Walt R. M.A. (Kansas) Asso. Tech. Adv., Div. Res. & Stat., Bur. Employ. Sec., Soc. Sec. Bd., Washington, D. C. *4104 N. Third Rd., Buckingham, Arlington, Va.*

Simon, Maj. Leslie E. Aberdeen Proving Gd., Md.

Singleton, Robert R. A.M. (Princeton) 20 Nassau St., Princeton, N. J.

Smart, Asso. Prof. L. Edwin Ph.D. (Ohio State) Ohio State University, Columbus, O. *410 King Ave.*

Smart, Oliver M. Ph.B. (Yale) Insp. Eng. Mat., U. S. Navy Bur. of Aeronautics, New York, N. Y. *156 Waverly Pl.*

Smith, John H. Ph.D. (Chicago) Instr., Univ. of Chicago, Chicago, Ill.

Snedecor, Prof. George W. M.S. (Michigan) Dir. Stat. Lab., Iowa State Coll., Ames, Iowa..

Soffer, Lewis R. B.S. (Chicago) Dir. of Res. & Stat., S-M News Co., 229 Fourth Ave., New York, N. Y.

Soloman, Herbert M.A. (Columbia) 801 E. Fremont Ave., Bronx, N. Y.

Soloman Herman S. B.S. (C. C. N. Y.) Stat. Super., N. Y. State Labor Dept., New York, N. Y. *1860 E. 12 St., Brooklyn, N. Y.*

South, Asso. Prof. Dudley E. Ph.D. (Michigan) Univ. of Kentucky, Lexington, Ky.

Spaney, Emma M.S. (C. C. N. Y.) Statistician, Co-op. Test Serv., Am. Coun. on Ed., New York, N. Y. *703 Chauncey St., Brooklyn, N. Y.*

Speeker, Asso. Prof. Guy G. A.M. (Indiana) Michigan State Coll., East Lansing, Mich. *Box 461.*

Spiegelman, Mortimer M.B.A. (Harvard) Metropolitan Life Ins. Co., One Madison Ave., New York, N. Y. *325 West 86 St.*

Stein, Arthur B.S. (C. C. N. Y.) Computer, War Dept., Aberdeen Proving Gd., Md. *Box 351, Aberdeen, Md.*

Stephan, Prof. Frederick F. M.A. (Chicago) Cornell Univ., Ithaca, N. Y. *102 McGraw Hall.*

Stewart, W. Mac Ph.M. (Wisconsin) 2230 Massachusetts Ave., NW, Washington, D. C.

Stevens, Milton S. M.A. (Columbia) 120 Kenilworth Pl., Brooklyn, N. Y.

Stouffer, Prof. Samuel A. Ph.D. (Chicago) Univ. of Chicago, Chicago, Ill.

Swanson, A. G. Ph.D. (Michigan) General Motors Inst., Flint, Mich.

Thom, Herbert C. S. Asso. Hydro. Eng., U. S. Weather Bur., Washington, D. C. *4767 21 Rd. North, Arlington, Va.*

Thompson, Dr. William R. 883 Warren St., Albany, N. Y.

Thomson, George W. B.S.E. (Michigan) Res. Chem., Ethyl Gasoline Corp., *1600 West 8 Mile Rd., Detroit, Mich.*

Thornton, Marian W. Ph.D. (Minnesota) Stat., Minn. Dept. of Health, Minneapolis, Minn. *852 Warren St., Albany, N. Y.*

Tiemann, Arthur E. A.B. (Harvard) Statistical Clerk, Bur. of the Census, Washington, D. C. *1405 Spring Rd.*

Tomlinson, John R. A.B. (Penna. State Coll.) Jr. Statistician, Aberdeen Proving Gd., Md. *Box 21, Aberdeen, Md.*

Toops, Prof. Herbert A. Ph.D. (Columbia) Ohio State Univ., Columbus, Ohio.

Toralballa, Leopoldo V. Ph.D. (Michigan) Asso. Actuary, Govt. Serv. Ins. System, Manila, P. I.

Torrey, Asso. Prof. Marian M. Ph.D. (Cornell) Asst. to Dean, Goucher Coll., Baltimore, Md.

Torrey, Mary N. 463 West St., New York, N. Y.

Treloar, Asso. Prof. Alan E. Ph.D. (Minnesota) Univ. of Minnesota, Minneapolis, Minn.

Tuttle, Charles R. M. Sc.D. (Nancy) Ord. Eng., War Dept., Washington, D. C. *726 31 St., SE.*

Upholt, William M. Ph.D. (California) Asst. Entomol., South Carolina Exp. Station, Clemson, S. C.

Van Voorhis, Walter R. Ph.D. (Penna. State Coll.) Instr., Fenn Coll., Cleveland, Ohio. *3147 Prospect Ave.*

Vatnsdal, Asst. Prof. John R. Ph.D. (Michigan) State Coll. of Washington, Pullman, Wash. *1916 B St.*

Vickery, C. W. Ph.D. (Texas) 310 Tribune Bldg., Austin, Tex.

Villavaso, Ernest J., Jr. B.A. (Texas) 3105 Duval St., Austin, Texas.

von Neumann, Prof. John Ph.D. (Budapest) Ing. Chem. (Zurich) Inst. for Advanced Study, Princeton, N. J.

Votaw, David F., Jr. M.A. (Texas) Asst., Princeton Univ., Princeton, N. J. *73 Graduate College.*

Wagner, Prof. Charles C. Ph.D. (Michigan) Adm. Asst. to Dean; Dir. of Lower Div., Penna. State Coll., State College, Pa.

Waite, Prof. Warren C. Ph.D. (Minnesota) Univ. of Minnesota. *Univ. Farm, St. Paul, Minn.*

Wald, Dr. Abraham Fayerweather Hall, Columbia Univ., New York, N. Y..

Walker, Harry A., Jr. Ph.D. (Rutgers) Asst., Squibb Inst. for Medical Res., New Brunswick, N. J.

Walker, Asso. Prof. Helen M. Ph.D. (Columbia) Teachers Coll., Columbia Univ., New York, N. Y.

Wallis, Asso. Prof. W. Allen B.A. (Minnesota) Stanford Univ., Calif.; Res. Asso., Natl. Bur. of Econ. Res., 1819 Broadway, New York, N. Y.

Wareham, Ralph E. B.A. (Iowa) Act. and Stat. Asst., General Electric Co., Schenectady, N. Y. *1123 Highland Park Rd.*

Watson, Alfred N. Ph.D. (Ohio State) Stat., Curtis Publishing Co., Philadelphia, Pa. *1408 Westwood Lane, Overbrook Hills, Pa.*

Webster, Redford T. Eng., Quality Control Dept., Western Electric Co., Kearney, N. J.

Wei, Dzung-shu Dept. of Math., St. John's Univ., Shanghai, China.

Weida, Prof. Frank M. Ph.D. (Iowa) George Washington Univ., Washington, D. C..

Weiner, Sidney B.S. (C. C. N. Y.) 1300 Monroe St., NW, Washington, D. C..

Welker, Dr. E. L. Univ. of Illinois, Urbana, Ill.

Welsch, Charles A. M.A. (New York) Instr., Adelphi Coll., Garden City, N. Y. *323 West 4 St., New York, N. Y.*

Wescott, Asst. Prof. Mason E. Ph.D. (Northwestern) Northwestern Univ., Evanston, Ill. *1936 Greenwood Ave., Wilmette, Ill.*

White, Prof. A. E. M.S. (Purdue) Kansas State Coll., Manhattan, Kan.

Wiesenberg, Albert B.S. (C. C. N. Y.) 80 Van Cortlandt Park South, New York, N. Y.

Wilkinson, Roger I. B.S. (Iowa State Coll.) Technical Staff, Bell Telephone Labs., 463 West St., New York, N. Y.
Wilks, Asso. Prof. Samuel S. Ph.D. (Iowa) Princeton Univ., Princeton, N. J..
Williams, John D. B.S. (Arizona) Box 47, Laveen, Ariz.
Wilson, Elizabeth W. Ph.D. (Radcliffe) One Waterhouse St., Cambridge, Mass.
Wilson, W. P. Box 1939, University, Ala.
Winsor, Charles P. Ph.D. (Harvard) Res. Asso., Princeton Univ., Princeton, N. J. *20 Nassau St.*
Wolfenden, Hugh H. Fel. Inst. Act. Consult. Act. & Stat., 182 Rosedale Hts. Dr., Toronto, 5, Ont., Can.
Wolfowitz, Jacob M.A. (Columbia) 25 Sherman Ave., Staten Island, N. Y.
Woodbury, Max A. M.S. (Michigan) Av. Cad., U. S. Army, 248 University St., Salt Lake City, Utah.
Wright, Prof. Sewall Ph.D. (Harvard) Univ. of Chicago, Chicago, Ill..
Wyckoff, John F. M.A. (Yale) Instr., Trinity Coll., Hartford, Conn.
Youden, William J. Ph.D. (Columbia) Phy. Chem., Boyce Thompson Inst. for Plant Res., 1086 N. Broadway, Yonkers, N. Y.
Young, Prof. Gale J. M.S. (Chicago) Head, Dept. of Math. and Phy., Olivet Coll., Olivet, Mich.
Young, Louis C. M.S. (Mass. Inst. of Tech.) Statistician, Westinghouse Elec. & Mfg. Co., 653 Page Blvd., Springfield, Mass.
Zeiger, Edward C. M.A. (Columbia) Asst. Sec. Head, Guardian Life Ins. Co., 50 Union Sq., New York, N. Y.
Zubin, Joseph Ph.D. (Columbia) Asso. Res. Psychol., N. Y. State Psychiatric Inst.; Instr., Columbia Univ., New York, N. Y. *722 West 168 St.*
Zuckerman, Samuel M.A. (Brooklyn) Asst. Computer, Aberdeen Proving Gd., Aberdeen, Md. *Box 351.*

GEOGRAPHICAL DISTRIBUTION OF MEMBERS

UNITED STATES

ALABAMA. (1)

UNIVERSITY. W. P. Wilson.

ARIZONA. (1)

LAVEEN. Williams.

CALIFORNIA. (14)

BERKELEY. Dantzig, Eudey, Gintzler, Neyman, Scott, Shephard.
DAVIS. Baker.
LA JOLLA. Gordon, McEwen.
LOS ANGELES. Alter, Hinds, Hoel.
SANTA MONICA. Marcuse.
STANFORD UNIVERSITY. Bacon.

COLORADO. (2)

FORT COLLINS. Clark, Guard.

CONNECTICUT. (11)

HARTFORD. Dorweiler, Elston, Keffer, Kurtz, Satterthwaite, Wyckoff.
MIDDLETOWN. H. E. Arnold, Camp.
NEW HAVEN. Bliss, Fisher, Ore.

DISTRICT OF COLUMBIA. (42)

WASHINGTON. Been, Blake, Brady, Brookner, Carlton, M. S. Cohen, G. J. Cox, Daly, Day, Deming, W. D. Evans, Flaherty, Frankel, Friedman, Girshick, Greville, Hansen, Harshbarger, Hendricks, Hermie, Hurwitz, Jacobs, M. B. Katz, Knudsen, Kullback, Madow, Nisselson, Norton, Osborne, O'Toole, Perlo, Pines, Rosander, Sandomire, Sasuly, Seeley, Siller, Stewart, Tiemann, Tuttle, Weida, Weiner.

GEORGIA. (2)

ATLANTA. Barnes, Pierce.

IDAHO. (1)

EAGLE. Bridger.

ILLINOIS. (23)

CHICAGO. Barrett, Bartky, Bronfenbrenner, Brooks, Ettinger, Householder, Kohl, Lange, Leavens, Livesay, Mansfield, Smith, Stouffer, Wright.
EVANSTON. Katzoff.
GALESBURG. Morton.
ROCK ISLAND. Cederberg.
URBANA. Boley, Crathorne, Gibson, Schwartz, Welker.
WILMETTE. Wescott.

INDIANA. (1)

WEST LAFAYETTE. Burr.

IOWA. (7)

AMES. Cochran, A. J. King, Snedecor.
IOWA CITY. A. T. Craig, Davies, Knowler, Rietz.

KANSAS. (2)

MANHATTAN. Fryer, White.

KENTUCKY. (2)

BOWLING GREEN. Milroy.
LEXINGTON. South.

LOUISIANA. (2)

NEW ORLEANS. Chapman.
UNIVERSITY. P. H. Anderson.

MARYLAND. (20)

ABERDEEN. Bechhofer, Bellinson, Gause, Norden, Reno, Tomlinson.
ABERDEEN PROVING GROUND. B. I. Hart, Kent, Shaw, Simon, Stein, Zuckerman.
BALTIMORE. Lukacs, Merrell, Reed, Sachs, M. M. Torrey.
BETHESDA. Dorn.
COLLEGE PARK. Lancaster.
HAVRE DE GRACE. Nekrassoff.

MASSACHUSETTS. (14)

BOSTON. K. J. Arnold, Mode, Roberts.
CAMBRIDGE. Bowker, Deemer, Freeman, Huntington, Kelley, Mises, Rulon, E. W. Wilson.
MELROSE. Bryan.
SPRINGFIELD. L. C. Young.
WORCESTER. O'Callahan.

MICHIGAN. (23)

ALLEGAN. Elkin.
ANN ARBOR. Carver, Cotterman, C. C. Craig, Dwyer, Fischer, Glover, Naugle, Nesbitt, Northam.
DETROIT. J. H. Levin, Pixley, Reitz, Thomson.
EAST LANSING. Baten, Blanche, Crowe, Dressel, Speeker.
FLINT. Swanson.
LANSING. Henry, L. Katz.
OLIVET. G. J. Young.

MINNESOTA. (12)

MINNEAPOLIS. W. L. Hart, D. Jackson, P. O. Johnson, Kozelka, I. Levin, Moore, Mudgett, Treloar.
ROCHESTER. Berkson, Gage, Miner.
ST. PAUL. Waite.

MISSISSIPPI. (1)

STATE COLLEGE. Ollivier.

MISSOURI. (3)

ST. LOUIS. Feldman, Regan, Rider.

MONTANA. (1)

BOZEMAN. Livers.

NEBRASKA. (1)

OMAHA. E. P. Coleman.

NEVADA. (1)

RENO. Carlson.

NEW JERSEY. (25)

DOVER, A. C. Cohen.
ESSEX FELLS. Olmstead.
HILLSIDE. Kiernan.
KEARNEY. Webster.
MADISON. Battin.
MOUNTAIN LAKES. Hotelling, Shewhart.
NEWARK. Neifeld.
NEW BRUNSWICK. H. A. Walker.
NUTLEY. Faust.
PRINCETON. T. W. Anderson, Dixon, Doob, Flood, P. J. McCarthy, Macphail, Mosteller, Scheffé, Singleton, von Neumann, Votaw, Wilks, Winsor.
UPPER MONTCLAIR. Hildebrandt.
WESTFIELD. Foster.

New York. (96)

Albany. Malzberg, Thompson, Thornton.
Arverne. Kury.
Astoria. Kantorovitz.
Baldwin. Grove.
Bayside. Perotti.
Bronx. Book, Dutka, Kwerel, Laderman, Muhlendorf, Ratkowitz, Seiden, H. Solomon.
Brooklyn. Griffin, Morrison, Peiser, Press, H. S. Solomon, Spaney, Stevens.
Canton. Bates.
Crown Point. Henderson.
Elmhurst. Petit.
Flushing. Sells.
Ithaca. Curtiss, Embody, Guttman, Stephan.
Little Neck. Robinson.
New York. Aroian, Arrow, Bachelor, Bernstein, Bonis, Boschan, G. W. Brown, R. H. Brown, Burgess, Bushey, Chalmers, Court, Di Salvatore, Dodge, Elveback, Fertig, Fry, Gumbel, Haines, Head, Heidingsfield, Hoy, Hsu, Jablon, Juran, F. G. King, Konijn, Le Leiko, Lorge, Lotka, Marschak, Molina, Norris, Paulsen, Payne, Peach, Peterson, Preinreich, Romig, Roodkowsky, Roos, Rubin, O. M. Smart, Soffer, Spiegelman, M. N. Torrey, Wald, H. M. Walker, Wallis, Welsh, Wiesenberg, Wilkinson, Zeiger, Zubin.
Poughkeepsie. Hopper.
Port Washington. Kimball.
Riverdale. Durand.
Rochester. Dunlap.
Rockville Centre. Berger.
Schenectady. Wareham.
Staten Island. Wolfowitz.
Troy. Lengyel.
Woodhaven. Munch.
Yonkers. Ashcroft, Youden.

New Mexico. (1)

Albuquerque. Larsen.

North Carolina. (6)

Chapel Hill. Hagood.
Durham. Greenwood.
Paw Creek. Elting.
Raleigh. R. L. Anderson, Cell, G. M. Cox.

Ohio. (11)

Cincinnati. Horst.
Cleveland. Cowan, Van Voorhis.
Columbus. L. E. Smart, Toops.
Lorain. Hand.
North Canton. Mummery, Schug.
Oxford. Pollard.
Toledo. Lemme.
Wellington. Ruger.

Oregon. (3)

Corvallis. Hammer.
Eugene. Kossack.
Salem. Olshen.

Pennsylvania. (24)

Aliquippa. Schrock.
Bethlehem. Passano.
Bryn Mawr. Geiringer.
Lewisburg. Benson, Richardson.
Millersville. Boyer.
New Kensington. Johner.
Oakmont. Petrie.
Overbrook Hills. Watson.
Philadelphia. Koopmans, Mauchly, Shohat.
Pittsburgh. Blackburn, Calkins, Elkins, Hebley, McDiarmid, Netzer, Niver, Olds, Savulak.
State College. Graves, E. Johnson, Wagner.

Philippine Islands. (3)

Manila. Jaramillo, Mills, Toralballa.

Rhode Island. (2)

Providence. Bennett, Feller.

South Carolina. (2)

Clemson. Upholt.
Columbia. J. B. Coleman.

Texas. (7)

Austin. Dodd, Mood, Vickery, Villavaso.
Dallas. Mouzon.
Lubbock. Michie.
Waco. Perry.

Utah. (1)

Salt Lake City. Woodbury.

VIRGINIA. (11)

ARLINGTON. Bingham, Caine, Schultz Shelton, Simmons, Thom.

DAHLGREN. Dresch.

LEXINGTON. Royston.

LYNCHBURG. Risley.

STAUNTON. Owen.

VIENNA. Brandt.

VERMONT. (1)

BENNINGTON. Lundberg.

WASHINGTON. (2)

PULLMAN. Vatnsdal.
WOODINVILLE. Anthony.

WISCONSIN. (7)

MADISON. Beal, Eisenhart, H. P. Evans, Fox, Ingraham, Ozanne.
MILWAUKEE. Kenney.

FOREIGN MEMBERS

ARGENTINA. (2)

BANFIELD. Acerboni.
BUENOS AIRES. Barral-Souto.

BRAZIL. (1)

RIO DE JANEIRO. Kingston.

CANADA. (6)

CHATHAM, ONTARIO. Beall.
EDMONTON, ALBERTA. Keeping.
KINGSTON, ONTARIO. Edgett.
TORONTO, ONTARIO. De Lury, R. W. B. Jackson, Wolfenden.

CHINA. (3)

SHANGHAI. Chang, Shen, Wei.

ENGLAND. (2)

ILFRACOMBE, DEVON. Perryman.
MANCHESTER. Ross.

IRELAND. (1)

CORK. M. D. McCarthy.

THE INSTITUTE OF MATHEMATICAL STATISTICS

(Organized September 12, 1935)

OFFICERS FOR 1941

President:

HAROLD HOTELLING, Columbia University, New York City.

Vice-Presidents:

A. T. CRAIG, University of Iowa, Iowa City, Iowa.
H. C. CARVER, University of Michigan, Ann Arbor, Michigan.

Secretary-Treasurer:

E. G. OLDS, Carnegie Institute of Technology, Pittsburgh, Pa.

The purpose of the Institute of Mathematical Statistics is to stimulate research in the mathematical theory of statistics and to promote coöperation between the field of pure research and the fields of application.

Membership dues including subscription to the ANNALS OF MATHEMATICAL STATISTICS are $5.00 per year. The dues and inquiries regarding membership in the Institute should be sent to the Secretary-Treasurer of the Institute.